UBIQUIBODIES: ENGINEERED UBIQUITIN LIGASES WITH UNNATURAL SUBSTRATE SPECIFICITY FOR TARGETED PROTEIN SILENCING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Alyse Danielle Portnoff

January 2014

UBIQUIBODIES: ENGINEERED UBIQUITIN LIGASES WITH UNNATURAL SUBSTRATE SPECIFICITY FOR TARGETED PROTEIN SILENCING

Alyse Danielle Portnoff, Ph. D.

Cornell University 2014

The ubiquitin-proteasome pathway (UPP) is the main route of protein degradation in eukaryotic cells and aids in regulation of cell cycle and cellular homeostasis. This robust pathway can also be utilized for reverse genetics to accelerate the degradation of otherwise stable cellular target proteins. In this work, we present a generalizable approach for protein knockdown by developing chimera proteins, called "ubiquibodies", which combine the activity of an E3 ubiquitin ligase with the affinity of designer binding proteins (DBPs). Specifically, we have utilized the modular E3 ubiquitin ligase CHIP and replaced its natural substrate-binding domain with antibody mimetic binding domains to create various ubiquibodies. Next, we optimized the chimeric construct expression in *E. coli* and purified uAbs to test their functional activity *in vitro*. Ubiquibodies were evaluated for both their target binding and subsequent target ubiquitination *in vitro*. This was further analyzed using mass spectroscopy to determine substrate ubiquitination sites and chain linkages. Within the eukaryotic cellular context, ubiquibodies were tested for their ability to specifically ubiquitinate and degrade their target proteins. Finally, preliminary work was performed using rational design to improve uAb E2 specificity, ensure flexibility for substrate binding and reduce autoubiquitination. From this foundation, we foresee the ubiquibody technology being a powerful tool to enable the dissection of protein function, including post-translational modifications, and the selective degradation of proteins that underlie human disease.

BIOGRAPHICAL SKETCH

Alyse Portnoff was born in Pittsburgh, PA to Allen and Angie Prince, and older sister Amy. She grew up in Overland Park, KS and attended St. Teresa's Academy in Kansas City, MO. After graduating in 2002, she enrolled at Carnegie Mellon University in the Department of Chemical Engineering. The summer after her sophomore year, she was privileged to acquire a summer research position at Stowers Medical Research Institute in the lab of Dr. Susan Abmayr. Working under the guidance of graduate student Kiran Kocherlakota, Alyse found her passion for research. From 2004-2005 she studied at École Polytechnique Fédéral de Lausanne in Switzerland. Upon returning to Carnegie Mellon for her senior year, she began researching in the lab of Professor Phil Campbell under the guidance of graduate student Eric Miller. After graduating with a B.S. in Chemical Engineering and Biomedical Engineering from Carnegie Mellon in 2006, Alyse married Samuel Portnoff. She then began her graduate degree at Cornell University in 2007 and joined the lab of Professor Matthew DeLisa in January 2008. During this time, she learned alongside many talented graduate students and post-doctoral researchers and under the excellent guidance of her adviser, Matt DeLisa. Alyse was awarded an NSF GRFP honorable mention in 2008, an NSF GK-12 Fellowship for 2009-2010 and an NIH CBI training grant for 2010-2012. In November 2011, Alyse and Sam had their first son, Isaiah Portnoff. Alyse completed her PhD in January 2014.

Dedicated to my husband Sam,

who rejoiced with me at every success no matter how small,

reminded me that negative results are rarely as bad as they first appear,

took care of our home and son more often than I can recount,

and without whom this would not have been possible.

ACKNOWLEDGMENTS

I must first recognize the critical role Matt DeLisa has played in guiding and encouraging the research efforts contained in this dissertation. Additionally, DeLisa group members have been a constant source of information, given valuable feedback and been a joy to work with throughout the years. Outside of the research lab, the New Life Presbyterian Church and Graduate Christian Fellowship have cared for my family in amazing ways. Finally, I must thank my family (extended and immediate) for their support and encouragement throughout this process.

This work was supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Institutes of Health Chemical Biology Training Grant T32 GM008500 (fellowship to A.D.P.) and the National Science Foundation GK-12 DGE-0841291 (fellowship to A.D.P.).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

^{10}Fn3	10^{th} fibronectin type 3 domain
A, Ala	Alanine
Abs	Absorbance
ACN	Acetonitrile
Ambic	Ammonium bicarbonate
AMP	Adenosine monophosphate
AR	Ankyrin repeat
ARM	Armadillo repeat
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
BAG	Bcl-2-associated athanogene
BHK21	Baby hamster kidney 21 cells
β-gal	β-galactosidase
βTrCP	β-transducin repeat-containing proteins
BLI	Bio-layer interferometry
BSA	Bovine serum albumin
C, Cys	Cysteine
CASA	Chaperone associated selective autophagy
CDR	Complimentary determining region
CHIP	Carboxyl terminus of Hsc70-interacting protein
CID	Collision induced dissociation
CMV	Cytomegalovirus
CO_2	Carbon dioxide
COS-7	*Cercopithecus aethiops* kidney 7 cells with SV40 large T antigen
C-terminal	Carboxyl terminus
D, Asp	Aspartic acid
Da	Dalton
DARPin	Designed ankyrin repeat protein
DBP	Designer binding protein
DDA	Data-dependent acquisition
DI	Deionized
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
DTT	DL-Dithiothreitol
DUB	Deubiquitinating enzyme
E, Glu	Glutamic acid
E. coli	*Escherichia coli*
E. coli	*Escherichia coli*
E1	Ubiquitin activating enzyme
E2	Ubiquitin conjugating enzyme
E3	Ubiquitin ligase
E4	Ubiquitin elongation enzyme
EDTA	Ethylenediaminetetraacetic acid

eGFP	Enhanced green fluorescent protein
EIA	Enzyme immunoassay
ELISA	Enzyme-linked immunosorbent assay
EMEM	Eagle's Minimum Essential Medium
ESI	Electrospray ionization
F, Phe	Phenylalanine
FA	Formic acid
FBS	Fetal bovine serum
FT	Fourier transform
Fwd	Forward
G, Gly	Glycine
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GCN4	General control nonrepressible 4
GFP	Green fluorescent protein
gpD	Bacteriophage capsid protein D
h	Hours
H, His	Histidine
H_2SO_4	Sulfuric acid
HA, Hag	Human influenza hemagglutinin epitope
HCD	High energy collision dissociation
HDAC6	Histone deacetylase 6
HECT	Homologous to E6-AP COOH terminus
HEK293T	Human embryonic kidney 293 cells with SV40 large T antigen
HF	High fidelity
HH	Helical hairpin
HPLC	High pressure liquid chromatography
HPV	Human papillomavirus
HRP	Horseradish peroxidase
Hsc/Hsp	Heat shock cognate/Heat shock protein
I, Ile	Isoleucine
iNOS	Inducible nitric oxide synthase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
JNK2	c-Jun N-terminal kinase 2
K, Lys	Lysine
KCl	Potassium chloride
K_D	Equilibrium dissociation constant
kDa	Kilodalton
k_{off}	Dissociation constant
k_{on}	Association constant
L, Leu	Leucine
LB	Luria-Bertani media
LC	Liquid chromatography
LD	Long distance
LRR	Leucine-rich repeat
LTQ	Linear trap quadrupole

M, Met	Methionine
MBP	Maltose binding protein
Met1	Methionine 1, N-terminal
$MgCl_2$	Magnesium chloride
$MgSO_4$-$7H_20$	Magnesium sulfate heptahydrate
MOPS	4-Morpholinepropanesulfonic acid
MS	Mass spectrometry
MW	Molecular weight
N, Asn	Asparagine
Na_2CO_3	Sodium carbonate
NaCl	Sodium chloride
NCBI	National Center for Biotechnology Information
NEL	Novel E3 ligase
Ni-NTA	Nickel nitrilotriacetic acid
NP40	Nonidet P-40 (IGEPAL CA-630)
N-terminal	Amine terminus
ONPG	*ortho*-Nitrophenyl-β-galactoside
OPD	*o*-Phenylenediamine dihydrochloride
P, Pro	Proline
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCA	Protein complementation assay
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDB	Protein data bank
PHD	Plant homeo domain
polyUb	Poly-ubiquitin chains
PPi	Pyrophosphate, $P_2O_7^{4-}$
pRB	Retinoblastoma protein
PTM	Post-translational modification
Q, Gln	Glutamine
R, Arg	Arginine
R4	scFv13-R4
Rev	Reverse
RING	Really interesting new gene
RNA	Ribonucleic acid
RPM	Rotations per minute
S, Ser	Serine
S. cerevisiae	*Saccharomyces cerevisiae*
SCF	Skp1, Cullin-1, F-box protein complex
scFv	Single-chain Fv (variable-fragment)
SDM	Standard deviation from the mean
SDS	Sodium dodecyl sulfate
SPR	Surface plasmon resonance
T, Thr	Threonine
T_m	Melting temperature

TPR	Tetratricopeptide repeat
TrisHCl	Tris(hydroxymethyl)aminomethane hydrochloride
uAb	Ubiquibody
Ub, Ubi	Ubiquitin
UBC	Ubiquitin conjugating domain
Ub_n	Ubiquitin chain of length n
UPP	Ubiquitin proteasome pathway
V, Val	Valine
V_H, V_L	Variable domain, heavy or light chain
W, Trp	Tryptophan
w:v	Weight to volume ratio
Y, Tyr	Tyrosine

CHAPTER 1

INTRODUCTION: HIJACKING THE UBIQUITIN PROTEASOME PATHWAY

Introduction

Harnessing the major native pathway of eukaryotic protein degradation has been a goal towards enabling cellular protein knockdown since the discovery of the ubiquitin-proteasome pathway [1-8]. Unlike reverse genetic tools that function at the DNA and RNA levels, targeted proteasomal degradation would allow for fast, highly specific, effective removal of proteins at the post-translational level. This attenuated removal of proteins is greatly desired for the amelioration of disease states in which detrimental proteins differ from their benign counterparts solely by post-translational modifications such as phosphorylation or glycosylation [9]. Additionally targeted proteolysis would allow for the continued study in differentiating protein functions pre- and post-translational modification.

Current reverse genetic methods for studying protein functions rely upon depletion of a protein by genetic knockout at the genomic level, or gene silencing by antisense deoxyoligonucleotides or RNA interference [10]. However, these methods remove all forms of the protein of interest regardless of its activity or post-translational state. Many protein functions are regulated intracellularly by post-translational modifications such as small chemical modifications (e.g. phosphorylation or glycosylation), protein modifications (e.g. ubiquitination or sumoylation) or proteolytic cleavage (e.g. pro-insulin cleavage). Thus, in order to study proteins at their functional level, post-translational technologies have been developed to inhibit proteins of interest. Major techniques include chemical inhibitors, high affinity antibody mimetic proteins and subcellular localization techniques (e.g. knock sideways); none of which actually remove the protein of interest from the cell [11, 12]. To address the desire for post-

translational knockout, several techniques have been developed which harness the ubiquitin proteasome pathway (UPP) for targeted protein degradation.

The ubiquitin proteasome pathway

Ubiquitination is a highly conserved post-translational modification involved in almost all cellular processes in eukaryotes, whereby proteins are modified by the addition of the small 76 amino acid protein appropriately named ubiquitin [13]. This modification can lead to sub-cellular protein trafficking, protein kinase activation, modulation of transcription factor activity, synthesis of DNA repair or proteasomal degradation [14]. Of these, ubiquitination leading to proteasomal degradation is the most characterized to date, as it was one of the earliest roles of ubiquitination identified [13].

The first step of the ubiquitination process occurs when the ubiquitin activating enzyme (E1) activates the C-terminal glycine of ubiquitin by ATP hydrolysis, then transfers the adenylated ubiquitin to the active site cysteine of the E1, yielding a reactive E1-ubiquitin thioester intermediate (**Fig. 1.1**) [15]. Upon the formation of an E1-E2 complex, the ubiquitin molecule is transferred to a conserved catalytic cysteine in the ubiquitin-conjugating (UBC) domain of the E2, ubiquitin conjugating enzyme [14]. Finally the ubiquitin ligase (E3) interacts with both the ubiquitin-charged E2 and the substrate protein and facilitates the transfer of ubiquitin from the E2 to the ε-amine group of a surface lysine residue on the substrate protein. Polyubiquitination (polyUb) or chain elongation occurs through a concerted effort of the E2 and E3, though the details of this process have not been fully elucidated [14-16].

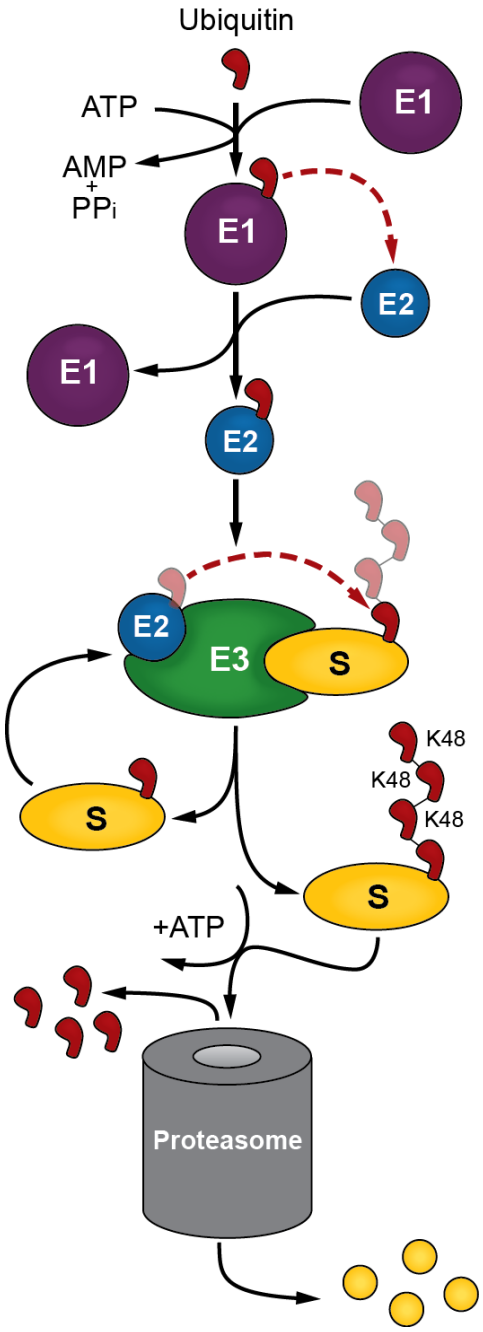


Figure 1.1 The ubiquitin proteasome pathway (UPP). Natural substrates (S) of the UPP are targeted for proteasomal degradation by the concerted activities of three enzymes. The sequential steps of ubiquitin (red) transfer are catalyzed by activating (E1), conjugating (E2), and ligase (E3) enzymes. For RING or U-box type E3s, the ubiquitin is transferred directly from the E2 to a surface lysine on the substrate protein (as depicted). Polyubiquitin chains are then formed on surface lysines of ubiquitin, which can direct for proteasomal degradation. The example shown is the canonical lysine-48 linked chain. Image adapted from Fang & Weisman [13].

In general, E3s are categorized into two main classes, the HECT (homologous to E6-AP COOH terminus) family and the RING (really interesting new gene) family. HECT ubiquitin ligases contain a conserved C-terminal domain (~350 amino acids) which accepts the ubiquitin from the E2 onto its own active site cysteine before the final transfer to the target protein. Contrastingly, RING ubiquitin ligases act solely as a bridge between the E2 and substrate protein (**Fig. 1.1**). These enzymes are classified by their RING domain made up of eight conserved cysteines and histidines that coordinate two zinc ions in a cross-braced manner [13]. Other classes of E3s related to the RING family include proteins with a PHD (Plant Homeo Domain), B-box, or U-box domain. For example, in the case of the U-box domain, the RING zinc-binding sites are replaced by charged and polar residues engaged in hydrogen-bonding networks required for structure and activity [17].

Increasing the complexity of ubiquitin modification is the possibility for multi-monoubiquitination or the formation of polyubiquitin (polyUb) chains; the compilation of which are referred to as the ubiquitin code (**Fig.1.2**). The protein ubiquitin contains seven lysine residues itself (K6, K11, K27, K29, K33, K48 and K63) and an N-terminal amine, all of which can function as receptor sites for ubiquitination. Though oftentimes the E2 or E3 is able to facilitate chain growth, sometimes the creation of polyUb chains requires a fourth enzyme called the ubiquitin elongation enzyme (E4) [15]. Specific chain formations differ in both structure and function and can cause various outcomes for the substrate protein, as shown in Figure 1.2 [14].

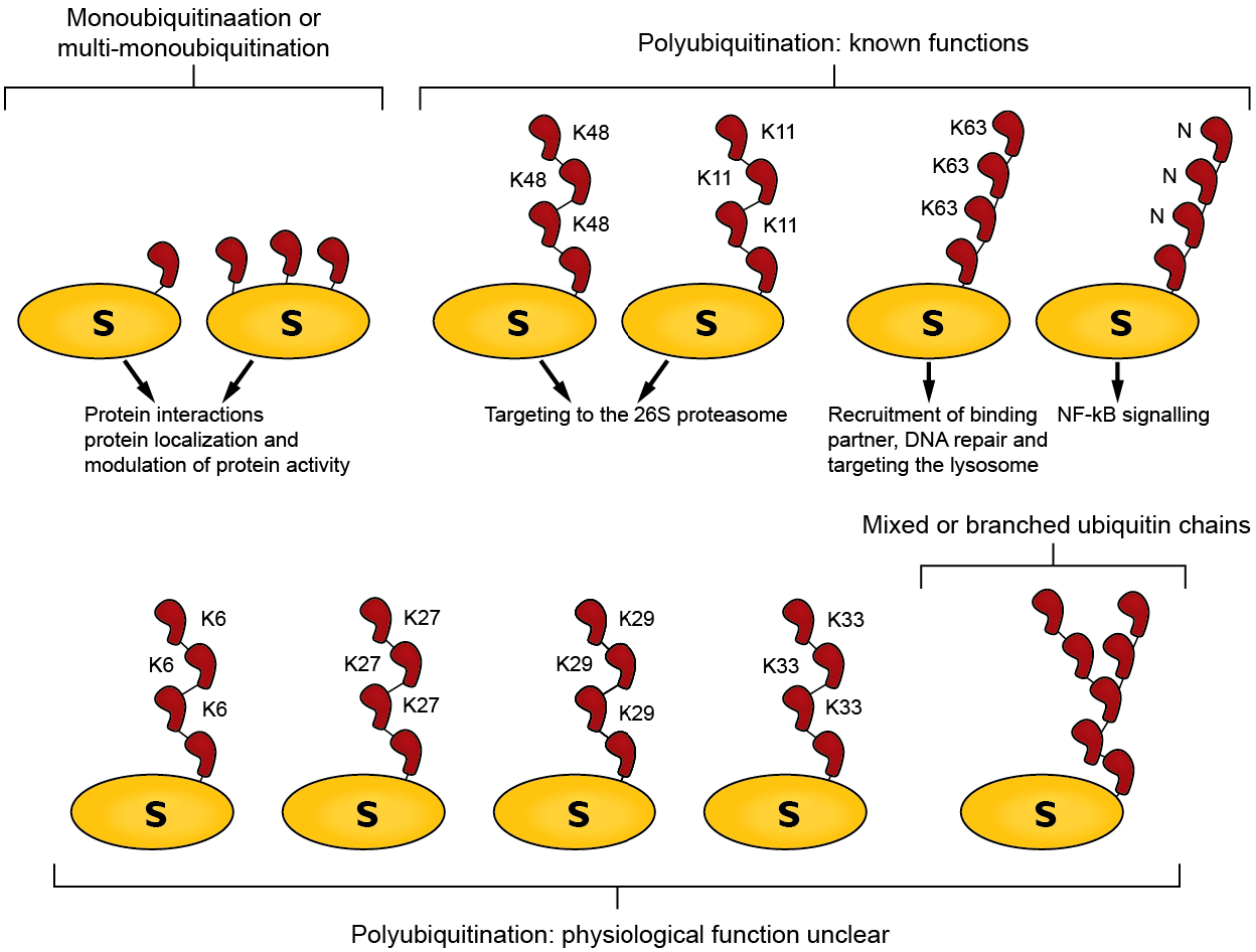


Figure 1.2 The ubiquitin code. Ubiquitin is attached to the ε-amino group of a lysine residue on substrate proteins (monoubiquitination). Further ubiquitin modification can occur by multi-monoubiquitination or the process of chain elongation. The different type of ubiquitin modification determines the cellular fate of the modified substrate, though not all ubiquitin chain motifs have well understood functions. Image adapted from Ye & Rape [14, 16]

Generally, the ubiquitin code is divided into two main classes, proteolytic and non-proteolytic functions [16]. Proteolytic functions include the regulation of proteasomal degradation and lysosomal degradation. For example, Lys48-linked chains, the most abundant linkage in all organisms, are the canonical chain linkage for proteasomal degradation [16]. However, Lys11-linked chains play the significant role of triggering degradation of cell cycle regulators during mitosis [18]. Additionally Lys29- and Lys63-linked chains have been implicated in proteasomal degradation in more specialized cases, revealing the flexibility of the

proteasomal cap proteins which recognize substrates for degradation [16]. Lysosomal degradation is generally mediated by monoubiquitination or Lys63-linked chains which target membrane proteins for lysosomes. Alternatively, ubiquitinated protein aggregates can be degraded by lysosomes after passing through autophagosomes, as directed by receptors recognizing Lys63-linked chains [19].

The well characterized non-proteolytic functions of ubiquitin modification are generally signaled by monoubiquitination, N-terminal or Lys63-linked chain formation [16]. Monoubiquitination has been shown to recruit protein binding partners and thereby mediate protein signaling (e.g. PCNA in DNA repair) [13]. Similarly Lys63-linked chains have been shown to regulate protein interactions, such as stabilizing polysomes and promoting translation [20]. Alternatively, ubiquitination can interfere with protein-protein interactions (including conformational changes) or activate a protein by targeting an inhibitory protein for degradation [16]. Finally, ubiquitination is known to regulate protein localization, such as the nuclear transport of multi-monoubiquitinated p53 [21]. Thus, while some of the ubiquitin code is well understood, the functions of many possible chain formations, including mixed or branched chains, are still being investigated [14, 16].

Harnessing the ubiquitin proteasome pathway

Since the discovery that a chain of four Lys48-linked ubiquitins could target proteins for proteasomal degradation, scientists have been trying to harness this regulated ubiquitin-proteasome pathway (UPP) for targeted protein silencing [22]. The earliest test of altering the UPP was performed *in vitro* by extending the C-terminus of ubiquitin conjugating enzymes (E2s) [1, 2]. The Vierstra group was evaluating the E2 protein domains that determined substrate specificity and decided to redirect the UPP by adding non-native protein binding domains to the

C-terminus of E2s and testing for the degradation of the binding partner *in vitro*. However, further studies of the UPP showed that ubiquitin ligases (E3) were the key step in protein specificity. Accordingly, the next manipulation of the UPP was with a ubiquitin ligase fusion created while studying the E6 and E7 proteins of oncogenic human papillomavirus (HPV) types 16 and 18 [7]. Scheffner et al. discovered they could degrade the retinoblastoma protein (pRB) *in vitro* by creating a fusion of the N-terminus of HPV-16 E7, which natively binds pRB, with the full-length HPV-16 E6, which natively binds and targets p53 for proteasomal degradation. The first successful *in vivo* redirection of ubiquitination for proteasomal degradation was performed by engineering the E3 multimeric protein complex SCF (Skp1, Cullin, F box-containing proteins) in yeast [8]. Zhou et al. showed that by utilizing the F box-containing protein Cdc4p, which is the substrate recognition component, and creating a chimeric protein with the known binding partner of a new target protein, that the entire SCF complex could be redirected toward degradation of the unnatural target protein. Furthermore they showed that the human homolog of Cdc4p, βTrCP could similarly be engineered and directed to selectively degrade the hypophosphorylated form of the target protein [8, 23-25].

Unfortunately, F-box subunits are thought to be constitutively unstable due to ubiquitination by the complex, for the specific purpose of enabling rapid remodeling of the SCF core [26]. Additionally, the over-expression of F-box chimeras was found to overload the core SCF complex and hindered ubiquitination of both native and novel target proteins [25]. A simpler E3 chimera protein was created by utilizing the U-box ubiquitin ligase CHIP (carboxyl terminus of Hsc70-interacting protein) and fusing it to a known binding partner of c-Myc, a proto-oncogene transcription factor [4]. This successful example of a redirected ubiquitin ligase proved the potential therapeutic use of protein knockdown in showing the suppression of tumor

formation by targeted degradation of c-Myc and was followed by a similar approach for targeted degradation of KRAS [27]. However in all of these cases, a pre-existing binding partner was required to facilitate the degradation of the targeted protein; thus for each new target, a novel engineered ubiquitin ligase would have to be created.

This shortcoming was recently addressed by Caussinus et al. who created a SCF ligase F-box protein chimera with a single-domain antibody fragment (i.e., nanobody) specific for green fluorescent protein (GFP), which was capable of depleting target GFP fusions [28]. While this new development will be widely applicable due to the availability of model organisms with genetic GFP fusions, it is unable to target endogenous proteins and unique isoforms of proteins. Alternative approaches in redirecting the UPP, which each have their own limitations, have linked multiple target proteins, utilized small synthetic ligands for specified binding or fused aptamers to an E3 [3, 5, 29]. However, a widely applicable, facile approach for effective protein knockdown to be used in reverse genetic studies has yet to be attained.

Designer binding proteins

The current forerunner in inhibiting protein function post-translationally has been the development of high affinity designer binding proteins (DBP). DBPs first developed out of technologies which utilized the natural diversity, specificity and affinity of antibodies from the immune system to label proteins of interest *ex vivo* for biochemical research purposes (e.g. immunoblotting, immunoprecipitation, cell sorting, etc.) However in order to utilize antibodies *in vivo,* dramatic changes were made to reduce their complexity and size and improve their intracellular stability.

One widely used antibody format is the single-chain Fv (variable-fragment) or scFv which is a fusion of the variable domain of the heavy chain (V_H) and the variable domain of the

light chain (V_L) connected by a flexible peptide linker to create a single-chain protein of about 28 kDa (**Fig. 1.3b**). This smaller format maintains the antigen binding site of the antibody which is harbored in the six complementary determining regions (CDRs), three each in the variable domains of the light and heavy chains. The CDRs are also known as hypervariable regions in that they retain the most diversity of the antibody (**Fig. 1.3c & d**). However both the V_H and V_L domains require intramolecular disulfide bonds which are readily destroyed in the intracellular reducing environment, thus limiting the applicability of scFvs intracellularly.

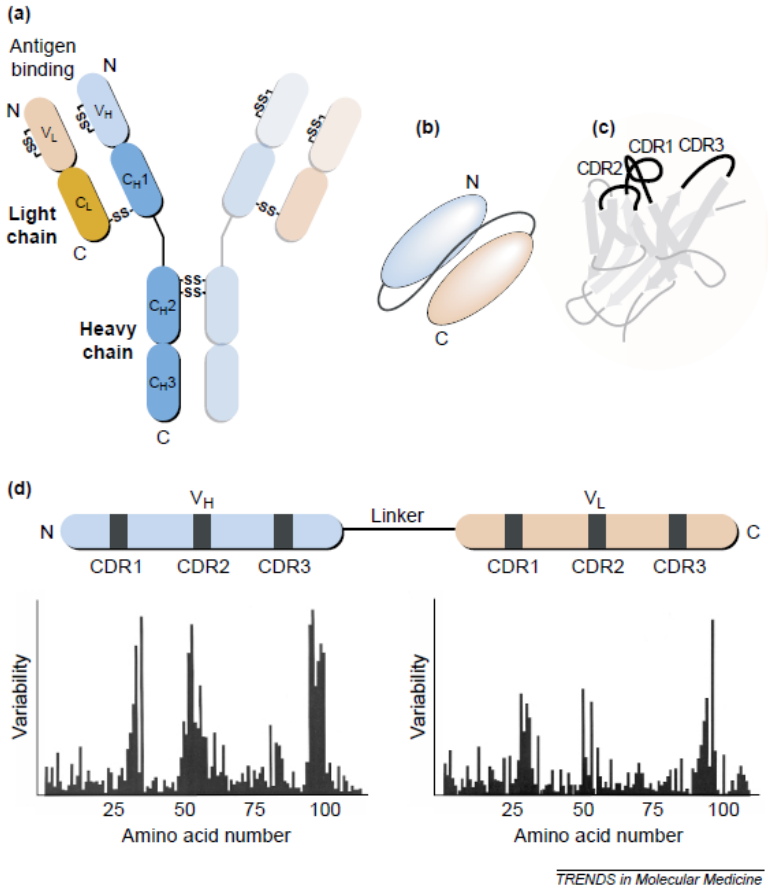


Figure 1.3 The scFv antibody fragment. (a) The antibody structure is a tetramer, made up of two light and two heavy chains connected by disulfide bonds (S-S). (b) The scFv antibody fragment is made by fusing the V_H and V_L domains using a flexible linker. (c) The antigen binding site of an antibody is made up of three CDR domains in each variable domain. (d) The diversity of binding in the antibody is made up of the hypervariable regions constituting the six CDRs. Thus the scFv format maintains the binding diversity of full-length antibodies. Image taken from Lobato & Rabbitts, [30].

To address the intracellular stability issue of scFvs, the concept of intracellular antibodies or intrabodies developed. Selection strategies for intrabodies isolate hits based on both high binding affinity and specificity, and the intracellular solubility. One method of identifying intrabodies is to add a second round intracellular screen post-phage display reduction of library size, such as intracellular antibody capture in yeast [31] or protein-activation in *E. coli* [32]. Alternative approaches utilizing the reducing cytoplasm of *E. coli* include intracellular ribosome display [33] and the hitchhiker mechanism of antigen-intrabody binding export through the twin-arginine translocation pathway [34]. Finally, various efforts have utilized the framework of established intrabodies as a scaffold for developing intrabody libraries [35-37].

In the process of creating smaller antibody formats with good affinity and stability, many synthetic biology tools were developed to enable the creation of gene libraries and selection techniques (e.g. phage display, ribosome display, surface display, protein complementation, etc.) [38]. However, with these synthetic technologies the use of the immunoglobulin scaffold itself became dispensable and thus protein engineers began to utilize novel binding proteins with the goal of improved biophysical properties (e.g. stability, solubility, multi-valency and modification capabilities) [39]. One major class of alternative scaffold proteins utilizes repeat protein domains such as ankyrin repeats (AR), leucine-rich repeats (LRR), tetratricopeptide repeats (TPR), HEAT repeats and armadillo repeats (ARM) [39]. Repeat protein domains are naturally found across various protein classes enabling protein-protein interactions and are characterized by small repeating structural motifs of 20-50 amino acids [39]. These structural motifs stack together to create the binding surface area of the protein, which enables binding specificity to evolve by both point mutations and insertion/deletion/shuffling of repeats [40]. As such, the modular repeat unit became an attractive scaffold for protein engineering. One such example is

the creation of designed ankyrin repeat proteins (DARPins) with varying numbers of internal ankyrin repeats (AR), each with six randomized residues, and capped with consensus N- and C-terminal ARs which provide stability (**Fig. 1.4c**) [39, 40]. Interestingly, it was later found that jawless vertebrates utilize LRRs to create diversity for their adaptive immune system [41], validating the use of repeat motifs for generating synthetic protein binding diversity [42].

Recognizing that nature uses many structural motifs for specific protein binding, alternatives to repeat-protein scaffolds have also been utilized to create synthetic libraries, such as monobodies, lipocalins and affibodies (**Fig. 1.4**) [43]. A monobody is based on the human 10^{th} fibronectin type 3 domain (^{10}Fn3) which is structurally similar to the β-sandwich structure of the V_H domain of an antibody. As such, these domains contain three loops, similar to the CDRs of antibodies, which have been diversified to create ^{10}Fn3 libraries (**Fig. 1.4a**) [44]. However the ^{10}Fn3 domain does not require disulfide bonds and is devoid of natural free cysteines, making it a candidate for both intracellular expression and site specific modification using cysteine chemistry. Another alternative, affibodies, were developed from the highly stable binding domain of *Staphylococcus aureus* protein A which has three α-helices and no disulfide bonds [42, 43]. Utilizing knowledge of the native binding interface, 13 surface exposed residues in the α-helices were randomized to create a combinatorial library from which novel binding proteins were selected [44]. As shown in Figure 1.4, this is divergent from the loop diversity found in monobodies and similarly scFv formats. Lastly, lipocalins have a rigid β-barrel with four flexible loops that create an entry to a ligand-binding cavity [44]. This scaffold has the unique advantage of selectively binding small molecules in the barrel/loop region. By randomizing the loop residues, novel binders have been isolated for both small molecules and protein-protein interactions [42]. Unfortunately, most natural lipocalins have intramolecular disulfide bonds

which aid in protein stabilization [43].

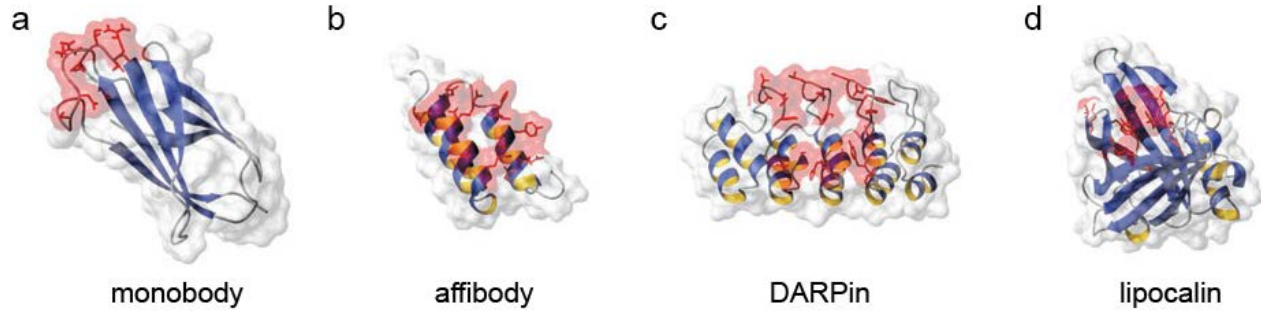


Figure 1.4 Alternative scaffold proteins. Each scaffold protein domain is shown as a ribbon diagram with residues randomized for library creation highlighted in red. (a) The monobody structure shows loop diversity similar to scFv domains, while (b) the affibody utilizes a flat surface for binding. (c) The DARPin repeat scaffold utilizes both surface and loop diversity in library creation, while (d) lipocalins have a recessed cavity for binding smaller molecules. Images taken from Binz et al. [42].

While the diversity of designer binding protein scaffolds each have unique advantages, all of the aforementioned motifs function by solely binding the protein of interest. As such, to inhibit the target protein the intracellular level of a DBP must be equal to, or exceed, that of the target because the target protein may escape the DBP binding. Furthermore, these DBPs may not fully neutralize the functional activity of target proteins, depending on their site specificity and size. Thus, we sought to develop a protein silencing strategy that links a DBP with the cell's natural degradation machinery – the ubiquitin proteasome pathway (UPP) – such that the steady-state levels of the target protein are systematically reduced.

Recent developments in targeted protein silencing

In this work we have developed a generalizable protein knockout method whereby an otherwise stable protein of interest is specifically targeted for proteasomal degradation. We have done this by engineering the final step of ubiquitination, namely the E3 ubiquitin ligase, and created chimeric proteins called "ubiquibodies" (uAb) which combine the versatile binding specificity of DBPs with the ubiquitination activity of a ubiquitin ligase to enable substrate

recognition and ubiquitination respectively (**Fig. 1.5**). Specifically, we have utilized the modular E3 ubiquitin ligase CHIP and replaced its natural substrate-binding domain with DBPs to create various ubiquibodies. Next, we optimized the fusion construct expression in *E. coli* and purified uAbs to test their functional activity *in vitro*. Ubiquibodies were evaluated for both their target binding and subsequent target ubiquitination *in vitro*. This was further analyzed using mass spectroscopy to determine substrate ubiquitination sites and chain linkages. Within the eukaryotic cellular context, ubiquibodies were tested for their ability to specifically ubiquitinate and degrade their target proteins. Furthermore, the modularity and generalizability of engineered uAbs were tested with multiple DBPs. Finally, preliminary work was performed using rational design to improve uAb E2 specificity, ensure flexibility for substrate binding and reduce autoubiquitination. From this foundation, we see the ubiquibody technology being a powerful tool to enable the dissection of protein function, including post-translational modifications, and the selective degradation of proteins that underlie human disease.

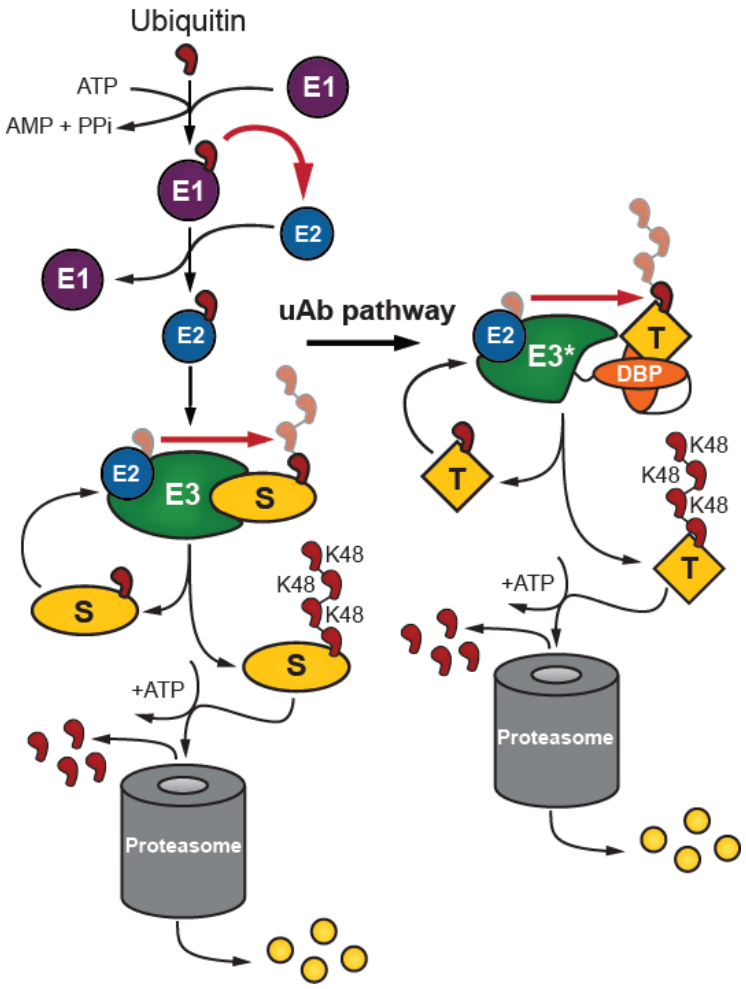


Figure 1.5 Redirecting the ubiquitin proteasome pathway. Schematic of redirecting the natural UPP with ubiquibodies (uAb). Naturally, the E1, E2 and E3 cascade tags substrate proteins (S) with polyubiquitin chains (K48 shown here) for proteasomal degradation. In the uAb pathway, the engineered E3*, where the natural substrate binding domain has been replaced with a designer binding protein (DBP), ubiquitinates the novel target (T) protein for proteasomal degradation.

Acknowledgements

I am greatly indebted to Matt DeLisa and Jeff Varner who first envisioned the idea of targeted protein degradation and gave me the freedom and direction to develop ubiquibodies. Additionally my knowledge of ubiquitination was enhanced by conversations with Sean O'Brien and by learning in the laboratory of Dr. Pengbo Zhou under the direction of Dr. Jianxuan Zhang.

CHAPTER 2

FUNCTIONAL EVALUATION OF AN ENGINEERED UBIQUITIN LIGASE

Introduction

In order to create an engineered ubiquitin ligase which could be redirected toward any cytosolic protein, we started with the soluble, modular E3 ubiquitin ligase CHIP (carboxyl terminus of Hsc70-interacting protein). CHIP is an ideal candidate E3 because it has well defined structural domains which are directly linked to functionality. Specifically, CHIP is composed of an N-terminal tetratricopeptide repeat (TPR) domain, a helical linker domain, and a C-terminal U-box ligase domain (**Fig. 2.1a**) [45]. The N-terminal TPR domain is involved in substrate recognition and is known to bind heat shock proteins such as Hsc70, Hsp70 and Hsp90 for which it was discovered and named [46]. As such, CHIP is involved in protein quality control by both aiding the folding of proteins in concert with the heat shock proteins or determining a substrate is unable to fold properly and thus marking it with ubiquitin for proteasomal degradation [47]. The helical linker domain of CHIP, also referred to as a helical hairpin or coiled-coil, contributes to protein flexibility which has been found to be essential for substrate ubiquitination [48]. Finally the C-terminal U-box domain binds E2 ubiquitin conjugating enzymes, enabling CHIP's E3 ubiquitin ligase activity. In addition to the defined structural domains, CHIP's natural substrate diversity in concert with heat shock proteins, suggested the possibility to further extend it towards non-native substrates for targeted proteolysis without disrupting its ubiquitin ligase activity.

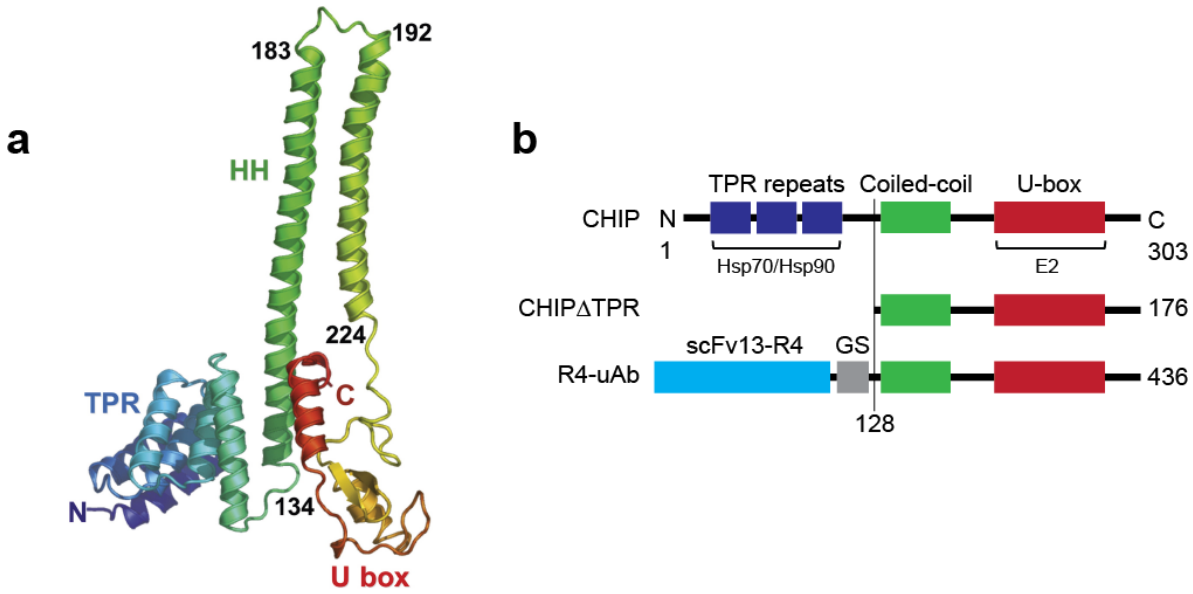


Figure 2.1 Engineering the E3 ubiquitin ligase CHIP. (a) The crystal structure of CHIP highlighting the functional domains from N- to C-terminus, TPR (tetratricopeptide repeat), HH (helical hairpin) and U-box domain taken from Zhang et al. [45]. (b) Linear representation of CHIP, CHIPΔTPR and R4-uAb. Numbers refer to amino acid positions from N-terminus (N) to C-terminus (C). The proteins are aligned vertically with the coiled-coil (a.k.a. HH) and U-box domains of CHIP. CHIPΔTPR is a truncated version of CHIP lacking the TPR domain. R4-uAb was designed with an additional Gly-Ser (GS) linker connecting the scFv13-R4 intrabody to CHIPΔTPR.

To redirect CHIP's ubiquitin ligase activity, we removed the TPR domain, its natural substrate binding domain, creating CHIPΔTPR (**Fig. 2.1b**). This domain was then replaced with a well characterized intrabody scFv13-R4 which binds β-galactosidase (β-gal) (**Fig. 2.1**). The intrabody scFv13-R4 was originally selected, after four rounds of mutagenesis, for its improved cytosolic expression in *E. coli* [32]. Thus our first antigen-ubiquibody pair for testing was β-galactosidase and the scFv13-R4-CHIPΔTPR fusion, hereafter referred to as R4-uAb (**Fig 2.1b**). This ubiquibody was first tested for soluble expression in *E. coli* and then purified to evaluate binding and ubiquitination activity *in vitro*. Furthermore, functionality of R4-uAb in the eukaryotic cellular context with the native UPP was tested to evaluate target ubiquitination and degradation efficiency.

Results

Soluble expression and purification of ubiquibodies in *E. coli.* Desiring high expression of uAbs for the purpose of purification, constructs were created in the medium copy plasmid pET28a (**Fig. 2.2a**) and expressed in the *E. coli* strain BL21(DE3) which is widely used for recombinant protein production [49]. Plasmid construction focused on building in modularity to enable facile swapping of DBP genes and thus each genetic unit was flanked with a unique restriction enzyme site (**Fig. 2.2a**). The addition of a Gly-Ser (GS) linker connecting the scFv13-R4 intrabody to CHIPΔTPR was also found to improve cytoplasmic expression (**Fig. 2.2b**). In order to mediate both 6xHis purification and facile immunoblotting (i.e. avoiding recognition of multiple 6xHis tagged proteins), a double epitope tag was tested at N- and C-terminal locations (**Fig 2.2b**). While tag locations had minimal affect on scFv13-R4 or CHIP alone, on the R4-uAb fusion, the C-terminal tag was significantly more stable (**Fig 2.2b**). Soluble expression of ubiquibodies compared to wild-type CHIP, CHIPΔTPR and scFv13-R4 showed an overall decrease in expression, but reveal that the fusion protein is stable and soluble (**Fig. 2.2c**). Furthermore, the expression of R4-uAb highlights the need for DBPs with intracellular stability as it can be compared to scFv13-uAb, a fusion with the parental scFv13 clone, which was not optimized for intracellular expression [32], and which could not be detected (**Fig. 2.2c**). Additional control constructs of R4-uAb include the point mutant R272A which is a U-box domain substitution known to inhibit E2 binding [50] and a non-specific uAb made with the intrabody D10 which binds the bacteriophage capsid protein gpD [51] (**Fig. 2.2c**). Each of these constructs was expressed and purified from the *E. coli* strain BL21(DE3) for *in vitro* functional evaluation (**Fig 2.2d**).

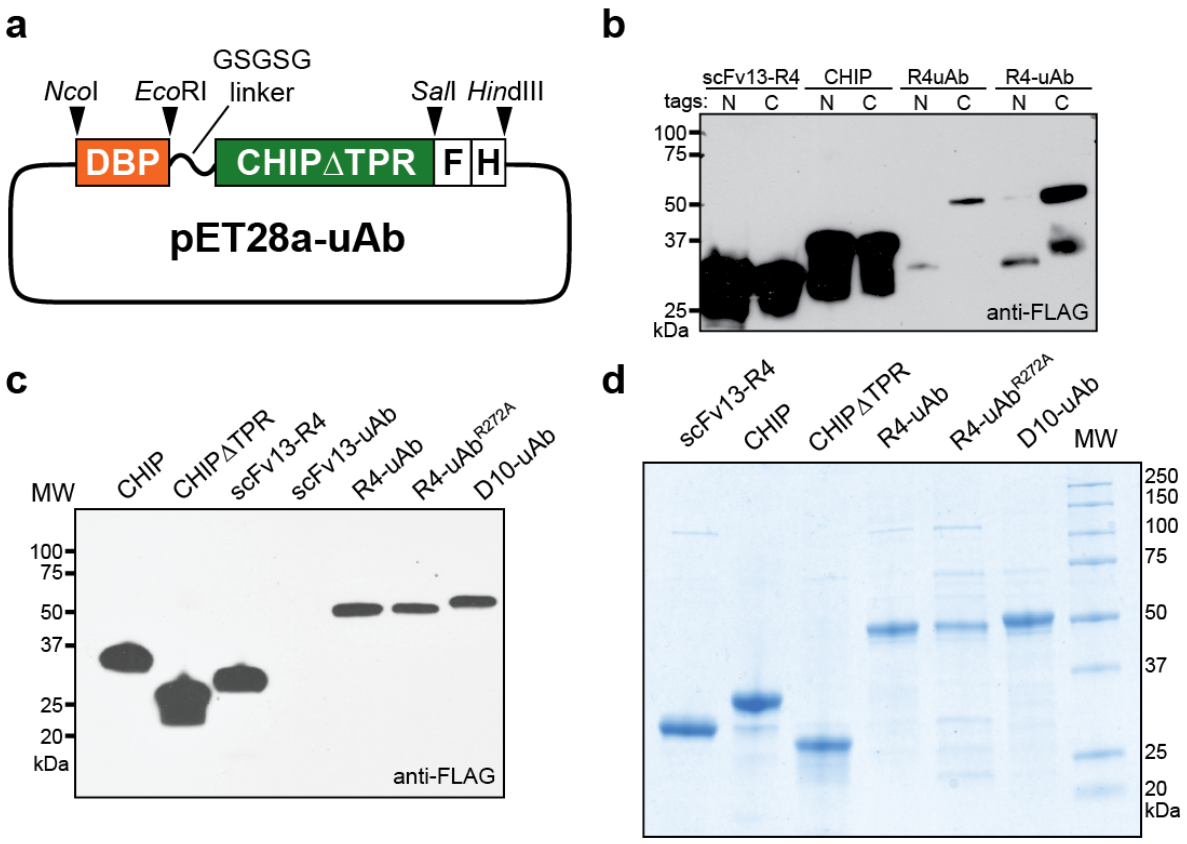


Figure 2.2 Expression of chimera proteins in *E. coli*. (a) Plasmid map of pET28a-uAb for expressing ubiquibodies in bacteria. DNA encoding designer binding proteins (DBPs) is cloned as a fusion with human CHIP that includes the coiled-coil and U-box domains but not the TPR domain (CHIPΔTPR). A flexible Gly-Ser linker (GSGSG) is introduced after the DBP. FLAG (F) and 6xHis (H) epitope tags are introduced after CHIPΔTPR. (b & c) Western blot analysis of cell lysates derived from *E. coli* strain BL21(DE3) expressing full-length scFv13-R4, CHIP, R4uAb direct fusion and R4-uAb fusion with the Gly-Ser linker (b). N- versus C-terminal double epitope tags are also denoted. (c) Comparison of different scFv uAb constructs made in pET28a-uAb (a). Specifically, parental scFv13-uAb is compared to intrabody R4-uAb. Also the control R272A point mutant which interferes with E2 binding and a non-specific scFv D10-uAb are included. An equivalent amount of total protein was loaded in each lane and anti-FLAG antibodies were used to detect the expressed proteins. (c) Coomassie stained SDS-PAGE analysis of different proteins from (b) following Ni^{2+}-affinity chromatography.

Functional testing of anti-β-gal ubiquibody *in vitro*. First we tested the binding affinity of purified R4-uAb compared to the scFv13-R4 alone and found that binding to β-gal was unaltered (**Fig. 2.3a**), indicating that fusion to CHIPΔTPR does not affect antigen-binding activity of the intrabody domain. Additionally the R272A substitution in the U-box domain had no affect on β-

gal affinity. In contrast, no significant β-gal binding was seen for CHIPΔTPR or the control D10-uAb (**Fig. 2.3a**).

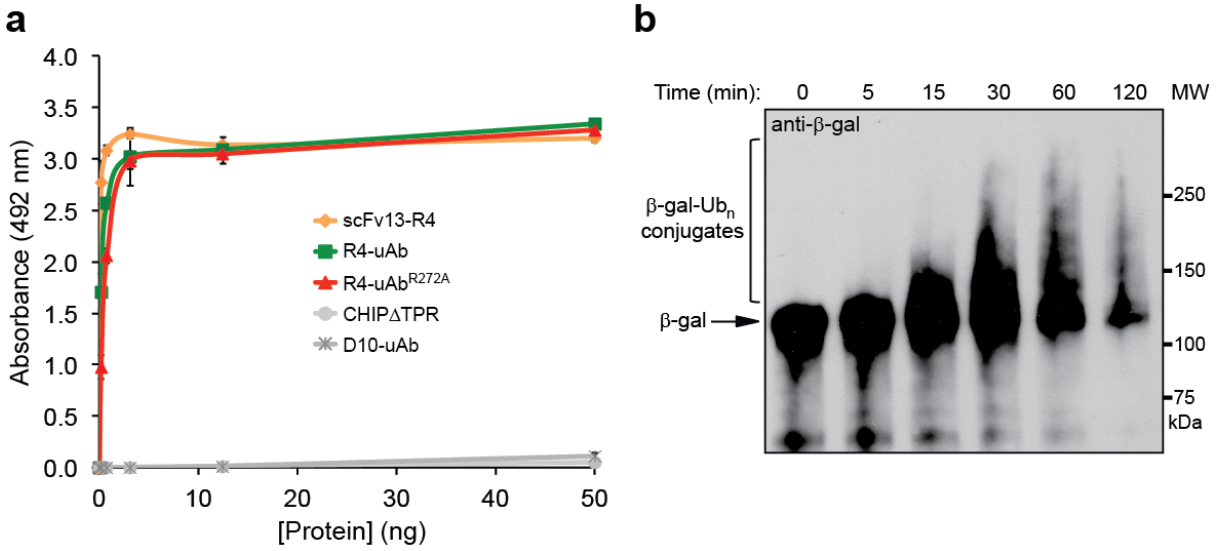


Figure 2.3 Redirecting a ubiquitin ligase towards a non-native substrate. (a) Binding activity of purified R4-uAb measured by ELISA with β-gal as antigen. The intrabody scFv13-R4 served as a positive control while CHIPΔTPR and D10-uAb served as negative controls. Also tested was R4-uAb^{R272A}, a derivative of R4-uAb carrying a point mutation in the U-box domain that is known to block the interaction between CHIP and the E2 enzyme, UbcH5α. (b) At the indicated times, *in vitro* ubiquitination reactions were stopped by boiling and immunoblotted with anti-β-gal antibodies. Protein bands corresponding to β-gal and the various β-gal-ubiquitin conjugates as well as the molecular weight of the marker bands (MW) are indicated. An equivalent amount of total protein was loaded in each lane.

Next, we performed *in vitro* ubiquitination assays with purified components, including R4-uAb as the E3 enzyme and β-gal as the target (note that β-gal has 20 lysine residues that serve as potential ubiquitin attachment sites [52]). UbcH5α was used as the E2 enzyme because it has previously been shown to function with CHIP *in vitro* [48]. High molecular weight bands corresponding to ubiquitinated β-gal were observed, especially over longer incubation times (**Fig. 2.3b**), which was characteristic of CHIP-mediated polyubiquitination of its native targets [48]. These results confirm that the CHIPΔTPR domain retained E3 ligase activity in the context of the chimera. Additionally, ubiquitination only proceeded when all pathway components were

included in the reaction (**Fig. 2.4**), indicating that R4-uAb activity was dependent on a complete ubiquitination pathway. Similar polyubiquitination was detected with an antibody specific for K48-linked polyUb chains (**Fig. 2.4**), confirming the presence of ubiquitin linkages that are known to signal proteasomal degradation [53]. Importantly, when reactions were performed with scFv13-R4, CHIPΔTPR, or the non-specific D10-uAb, there was no detectable ubiquitination of β-gal (**Fig. 2.4**). Likewise, R4-uAb^{R272A} which should not interact with the E2 UbcH5α, failed to conjugate ubiquitin (**Fig. 2.4**) even though it was capable of binding β-gal (**Fig. 2.3a**).

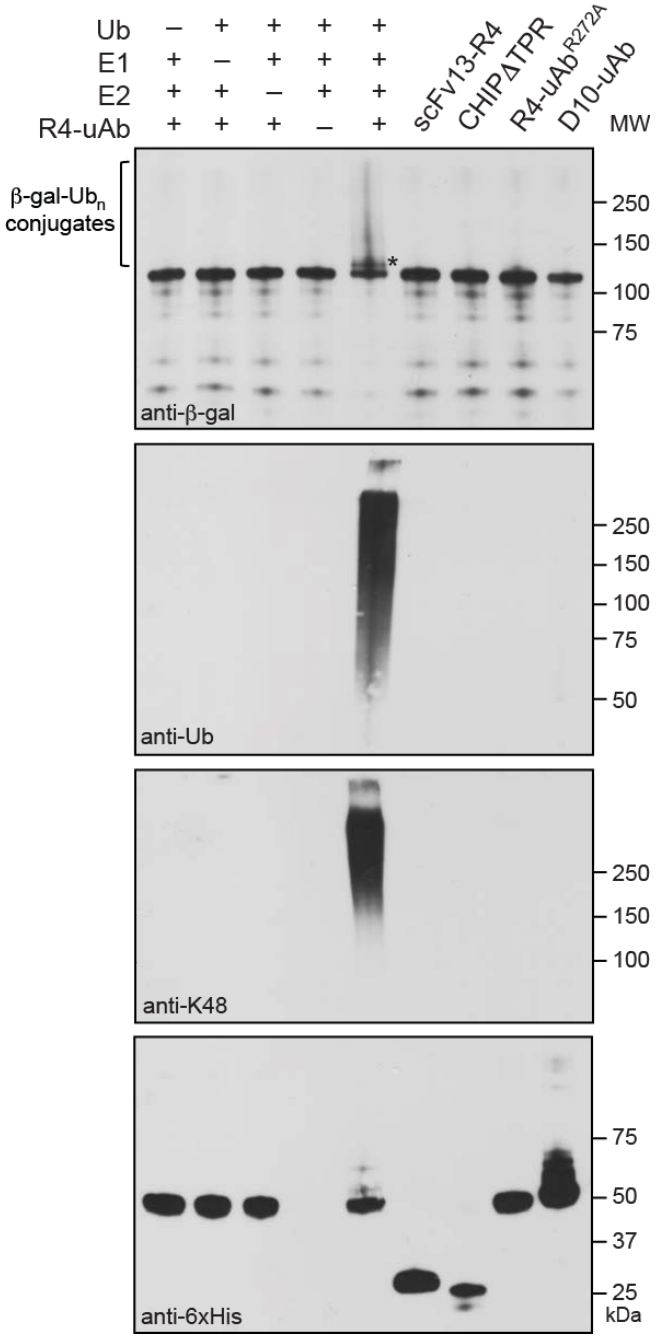


Figure 2.4 ***In vitro*** **ubiquitination of β-gal.** Ubiquitination of β-gal was evaluated in the presence (+) or absence (-) of each ubiquitin pathway component, namely ubiquitin (Ub), E1, E2, and R4-uAb as E3. Controls included scFv13-R4, CHIPΔTPR, R4-uAb^{R272A}, and D10-uAb. An equivalent amount of total protein was added to each lane. Immunoblots were probed with anti-β-gal, anti-ubiquitin, anti-K48, and anti-6xHis antibodies. Protein bands corresponding to β-gal, mono-ubiquitinated β-gal (*) and β-gal-ubiquitin conjugates as well as the molecular weight of the marker bands (MW) are indicated. The results are representative of at least three replicate experiments.

Identification of lysine residues modified by anti-β-gal ubiquibody. We next investigated which lysine residues were ubiquitinated by our engineered R4-uAb fusion. *In vitro* ubiquitination of β-gal was monitored by SDS-PAGE and the formation of higher molecular weight species was clearly evident in the region of the gel where ubiquitinated β-gal would be expected to resolve (**Fig. 2.5**). We also detected R4-uAb-ubiquitin conjugates, consistent with earlier reports showing autoubiquitination of CHIP and CHIP fusions [48]. However at early time points, these were at lower molecular weight regions of the gel. Therefore, bands on the gel corresponding to modified β-gal after 15 min of ubiquitination (delineated in red), were excised, digested with trypsin, and analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (**Fig. 2.5**).

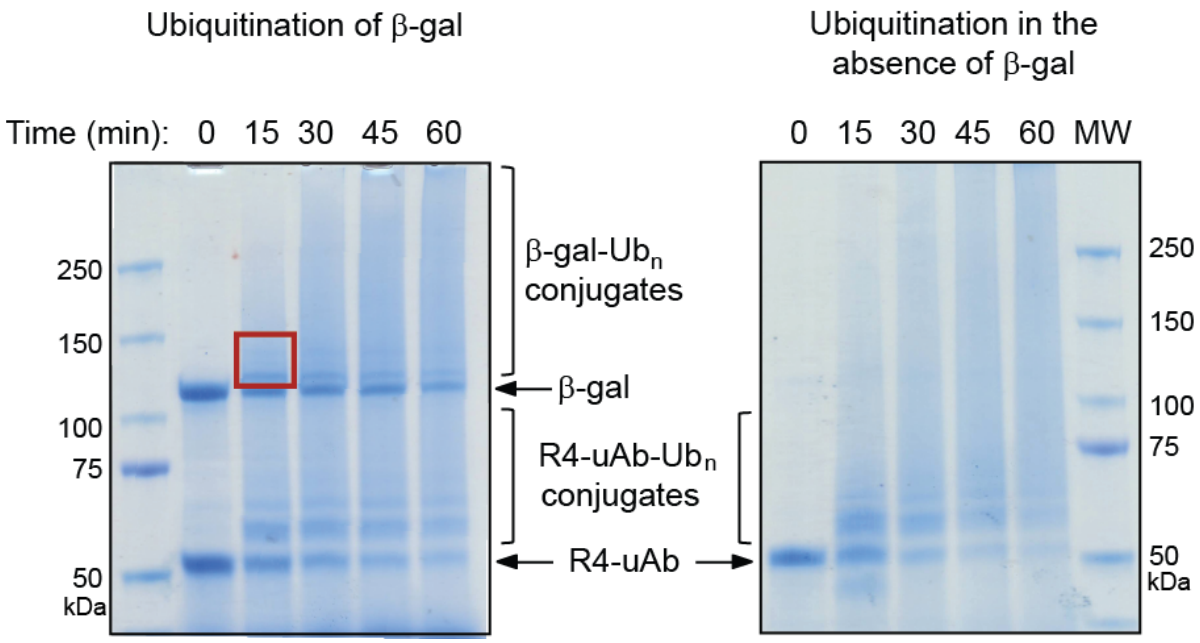


Figure 2.5 ***In vitro*** **ubiquitination by engineered ubiquibodies.** Coomassie-stained SDS-PAGE analysis of R4-uAb-mediated ubiquitination reactions in the presence and absence of *E. coli* β-gal at various times after initiation. Protein bands corresponding to unmodified β-gal, R4-uAb, various ubiquitin conjugates and the molecular weight of the marker bands (MW) are indicated. Trypsin digests and subsequent LC-MS/MS analysis was performed on the proteins delineated by the red box.

Peptides corresponding to 80% of the β-gal sequence were identified using Mascot software, suggesting good coverage of the entire protein sequence. Trypsin digestion of a ubiquitinated protein leaves the C-terminal Gly-Gly of ubiquitin attached to the ubiquitinated lysine residue. Therefore, the MS data was searched for such modification of β-gal lysines and five modified residues were identified: K348, K518, K662, K774, and K775. All five lysines are solvent accessible [52] (**Fig. 2.6a**), consistent with the location of ubiquitin attachment sites on native substrates Hsp70 and Hsp90 [54]. The MS/MS spectra of two β-gal peptides that include the identified ubiquitination sites K774 and K775 are depicted in Figures 2.6b and c, respectively.

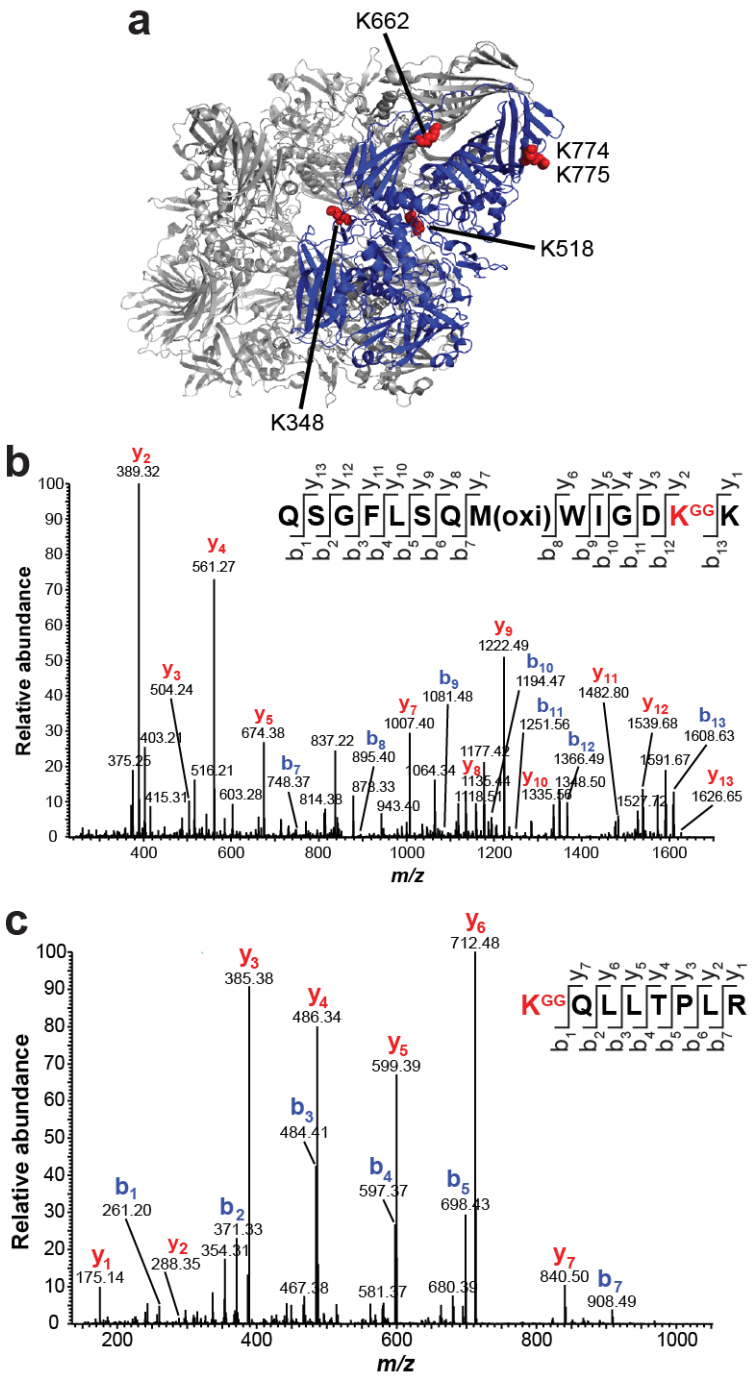


Figure 2.6 Lysine residues of β-gal modified by ubiquitin. (a) Mapping of ubiquitinated lysines (red balls) onto crystal structure of a single β-gal monomer (blue). All identified lysines are solvent exposed in the homotetramer (grey) model generated in PyMOL™ using PDB 1DP0. MS/MS spectra of representative β-gal peptides (b) 762-QSGFLSQMWIGDKK-775 and (c) 775-KQLLTPLR-782, containing Gly-Gly modified (ubiquitinated) lysine residues. Modified lysine residues are labeled with GG and correspond to (a) K774 and (b) K775 in β-gal. Fractionation of the peptides into b and y ions was performed, and the corresponding peaks are labeled on the spectra. The y-axis, relative abundance, was normalized to the most abundant identified peptide fragment.

Peptides corresponding to 95% of the ubiquitin sequence were also identified in the mass spectrometry analysis. Specifically, lysine residues K6, K11, K48, and K63 within ubiquitin were found to be modified in peptides isolated from our high molecular weight samples, suggesting these chain linkages in the polyubiquitination of β-gal by R4-uAb. These same chain linkages were previously observed on natural substrates that had been ubiquitinated by full-length CHIP *in vitro* [54]. The MS/MS spectra of two ubiquitin peptides that include the identified ubiquitination sites K48 and K63 are depicted in Figure 2.7a and b, respectively. It is worth noting that while K48-linked chains are considered the canonical linkage associated with targeting proteins for proteasomal degradation, all linkages identified here including K6, K11, and K63 have been implicated as targeting signals for the 26S proteasome [55, 56].

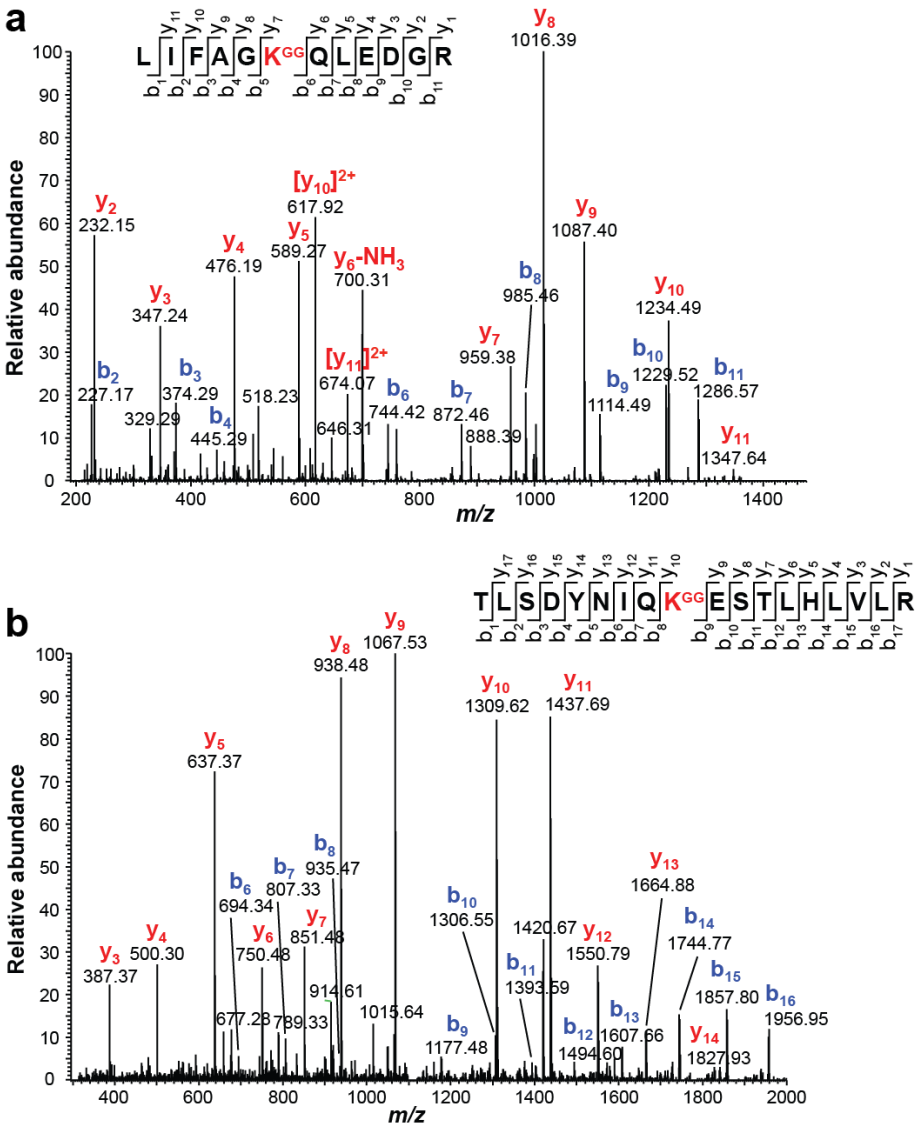


Figure 2.7 Ubiquitin lysine residues modified by another ubiquitin. MS/MS spectra of representative ubiquitin peptides (a) 43-LIFAGKQLEDGR-54 and (b) 55-TLSDYNIQKESTIHLVLR-72 containing Gly-Gly modified (ubiquitinated) lysine residues. Modified lysine residues are labeled with GG and correspond to (a) K48 and (b) K63 in ubiquitin. Fractionation of the peptides into b and y ions was performed, and the corresponding peaks are labeled on the spectra. Peptide fragment ions that have a +2 charge or lost an NH_3 group are indicated with 2+ and -NH_3, respectively. The y-axis, relative abundance, was normalized to the most abundant identified peptide fragment.

Ectopic co-expression of β-gal and R4-uAb. We next investigated whether R4-uAb-mediated ubiquitination would result in β-gal depletion by the UPP in mammalian cells. First, HEK293T cells were transiently co-transfected with pcDNA3-based plasmids encoding *E. coli* β-gal and the

R4-uAb, where each construct was under control of the strong human cytomegalovirus (CMV) promoter. Next, cellular β-gal levels were measured by immunoblotting cell lysates collected 24 h post-transfection. When HEK293T cells were co-transfected with pcDNA3-β-gal and increasing amounts of pcDNA3-R4-uAb, the β-gal levels were systematically reduced to as low as 2% of the steady-state levels measured in cells transfected with only the pcDNA3-β-gal plasmid (**Fig. 2.8**). In contrast, no reduction in β-gal expression was observed following co-transfection with the pcDNA3-D10-uAb or pcDNA3-scFv13-R4 plasmids (**Fig. 2.8**). Interestingly, cells co-transfected with R4-uAb^{R272A} exhibited an intermediate level of β-gal expression indicating that this point mutation in CHIP's U-box domain may inhibit some but not all E2 interactions *in vivo*. Importantly, the levels of a housekeeping protein, GAPDH, and a native binding partner of full-length CHIP, Hsp70, were not affected by co-transfections (**Fig. 2.8**).

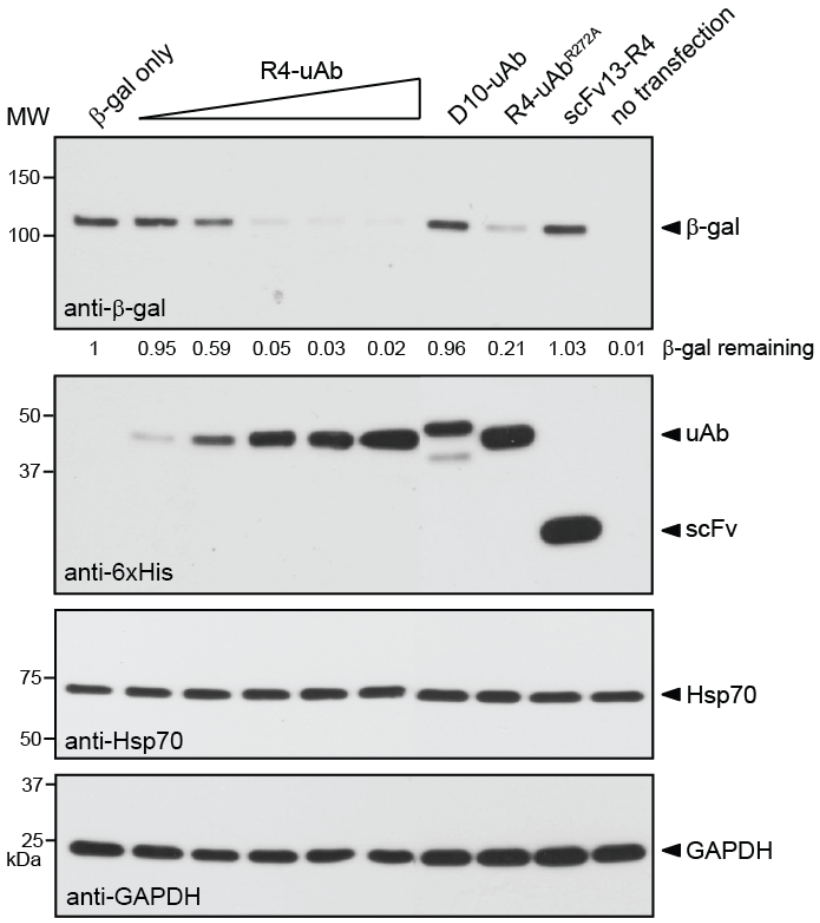


Figure 2.8 R4-uAb-mediated proteolysis of β-gal in mammalian cells. Immunoblots of extracts prepared from HEK293T cells transfected with pcDNA3-β-gal alone (β-gal only) or co-transfected with pcDNA3-β-gal along with one of the following: pcDNA3-R4-uAb, pcDNA3-D10-uAb, pcDNA3-R4-uAb^{R272A}, or pcDNA3-scFv13-R4. The triangle indicates increasing amounts of pcDNA3-R4-uAb plasmid DNA used to transfect cells. The percentages of β-gal remaining in each sample were quantitated by densitometry scanning and are indicated. An equivalent amount of total protein was loaded in each lane. Blots were probed with antibodies specific for β-gal, 6xHis, GAPDH and Hsp70 as indicated. The immunoblot results are representative of at least three replicate experiments.

Evaluating β-gal knockdown by microscopy. Next we began evaluating co-transfected HEK293T cells with microscopy techniques to further investigate the β-gal knockdown shown by immunoblotting. A pcDNA3-β-gal-eGFP fusion was created in order to perform live cell imaging at various time points post-transfection. At 8 h post-transfection, β-gal-eGFP expression was seen diffuse throughout the cellular cytoplasm (**Fig. 2.9** β-gal only). However,

when pcDNA3-β-gal-eGFP was co-transfected with pcDNA3-R4-uAb, fluorescent foci were seen which were not detected in control co-transfections with pcDNA3-scFv13-R4 nor pcDNA3-D10-uAb (**Fig. 2.9**).

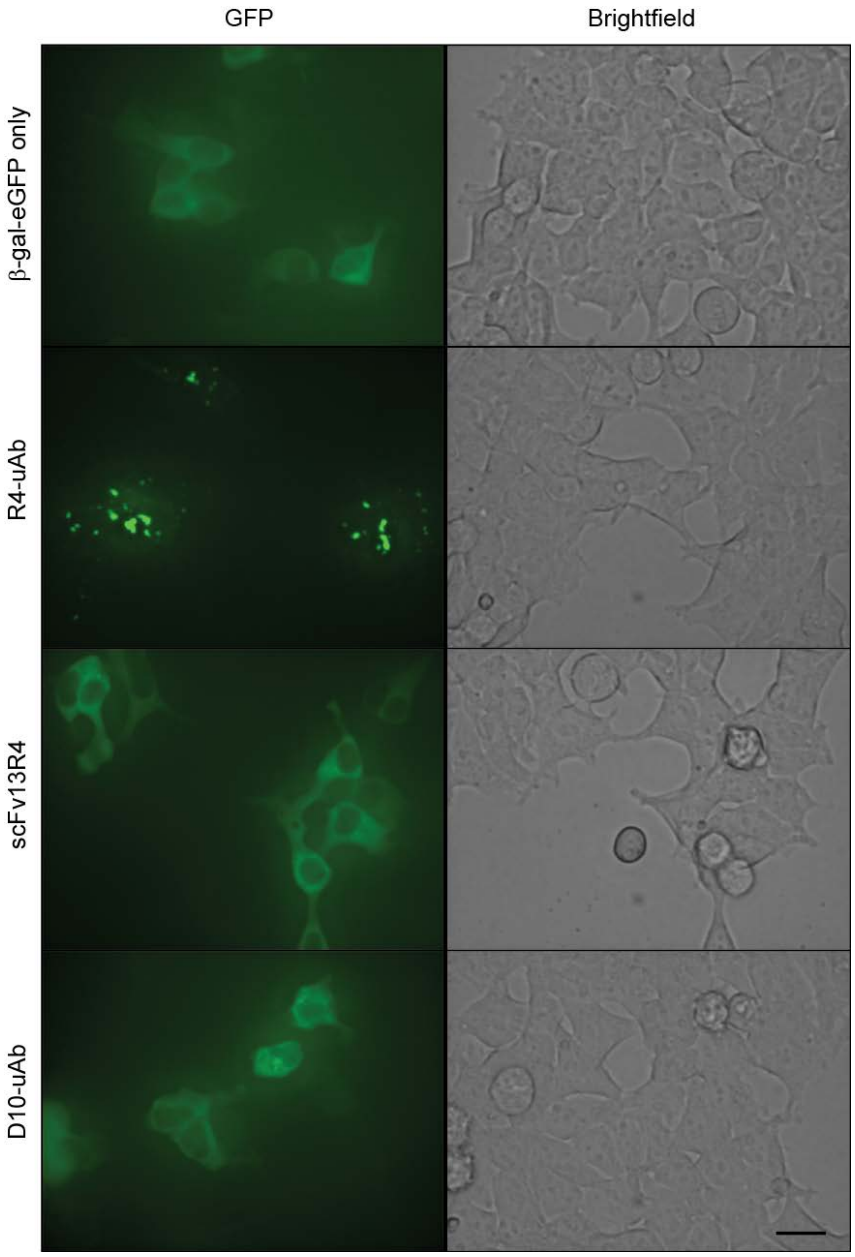


Figure 2.9 Fluorescence microscopy of β-gal-eGFP fusion. HEK293T cells were transfected with 0.5 μg pcDNA3-β-gal-eGFP alone (β-gal-eGFP only) or co-transfected with 1.25 μg pcDNA3-R4-uAb, pcDNA3-scFv13-R4 or pcDNA3-D10-uAb as indicated. Live cell images were taken at 8 h post-transfection using a 40x objective. The scale bar shown is 20 μm.

To determine if foci were an artifact of the β-gal-eGFP fusion, β-gal expression was next evaluated using immunofluorescent staining at 24 h post-transfection (**Fig. 2.10**). Again, when HEK293T cells were transfected with pcDNA3-β-gal alone, β-gal expression was diffuse throughout the cytoplasm of the cells, but when co-transfected with the pcDNA3-R4-uAb, foci were detected (**Fig. 2.10**).

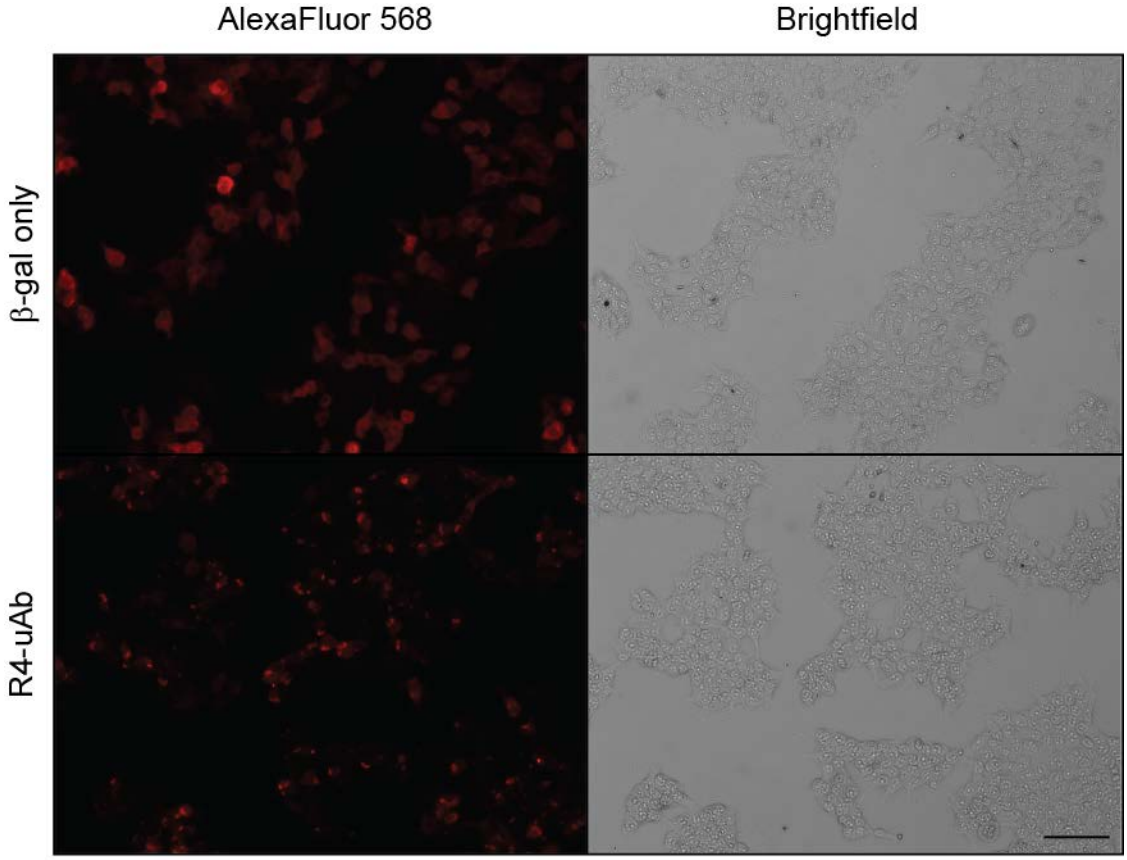


Figure 2.10 Immunofluorescent staining of 293T cells transfected with β-gal. HEK293T cells transfected with 0.5 μg pcDNA3-β-gal alone or co-transfected with 1.0 μg pcDNA3-R4-uAb as indicated. Cells were fixed at 24 h post-transfection, permeabilized and stained with anti-β-gal primary antibodies followed by a secondary conjugated with Alexa Fluor 568. Images were taken using a 5x objective and the scale bar shown is 40 μm.

One possible explanation for the foci was that β-gal was being partitioned to the insoluble fraction. To evaluate this possibility, HEK293T cells were transfected as in Figure 2.8 except that insoluble fractions were then solubilized in 2% SDS by boiling and loaded equivolume with

the soluble lysates for detection by immunoblotting (**Fig. 2.11**). Cellular fractionations revealed that both β-gal and β-gal-eGFP were partitioning to the insoluble fractions when cells were co-transfected with the ubiquibody (**Fig. 2.11**). Additionally the ubiquibody itself was found partitioning to the insoluble fraction (anti-6xHis). For these analyses, GAPDH served as a loading control for the soluble fractions and ensured that the insoluble fractions did not contain appreciable amounts of soluble cytosolic proteins (**Fig. 2.11**).

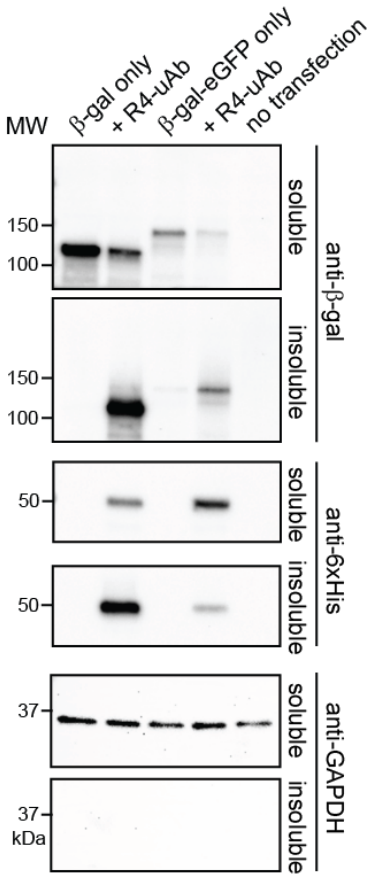


Figure 2.11 Evaluating insoluble fractions from 293T cells. Immunoblots of soluble and insoluble extracts prepared from HEK293T cells transfected with pcDNA3-β-gal alone (β-gal only), pcDNA3-β-gal-eGFP alone (β-gal-eGFP only), or co-transfected with pcDNA3-R4-uAb. An equivalent amount of total protein was loaded in each soluble lysate lane and equivolume samples were loaded for the insoluble fractions. Blots were probed with antibodies specific for β-gal, 6xHis and GAPDH as indicated and molecular weight (MW) markers are labeled.

Results from the insoluble fractionation revealed that the reduction in β-gal levels may be a combined effect of UPP degradation and partitioning to the insoluble fraction. In order to

investigate this phenomenon, we sought to determine if the partitioning was due to the high expression level of ubiquitinated β-gal, co-expression with an exogenous protein (i.e. non-specific to ubiquibody interactions), or over-expression of the specific uAb binding to β-gal. To address each of these theories, HEK293T cells were transfected with 5-fold (**Fig. 2.12a**) and 10-fold (**Fig. 2.12b**) less pcDNA3-β-gal and a broad range gradient of pcDNA3-R4-uAb DNA (**Fig. 2.12**). From these gradients, it was evident that at lower levels of transfection, less β-gal and less R4-uAb partitioned to the insoluble fractions. Furthermore, by reducing the pcDNA3-β-gal transfection 10-fold, partitioning of β-gal to the insoluble fraction was only detectable with prolonged exposure and R4-uAb was undetectable. Additionally, the controls pcDNA3-scFv13-R4 and pcDNA3-D10-uAb were included at the highest level of co-transfection and revealed no partitioning of β-gal to the insoluble fraction, revealing that the phenomenon was dependant on cognate ubiquibody co-transfection (**Fig. 2.12a** and **b**, right hand lanes).

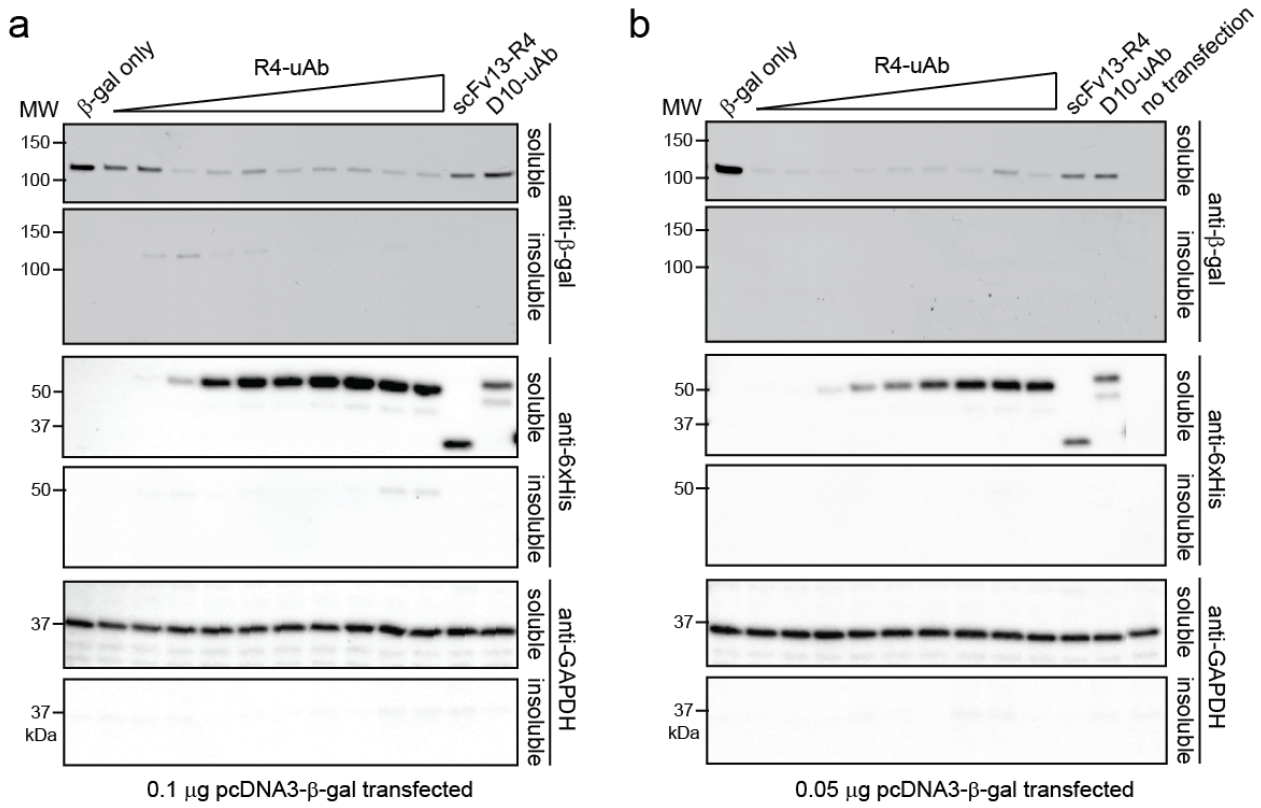


Figure 2.12 Reducing co-transfection levels of β-gal and R4-uAb. Immunoblots of extracts prepared from HEK293T cells transfected with pcDNA3-β-gal alone (β-gal only) or co-transfected with pcDNA3-β-gal and one of the following: pcDNA3-R4-uAb, pcDNA3-D10-uAb or pcDNA3-scFv13-R4. The triangle indicates increasing amounts of pcDNA3-R4-uAb plasmid DNA used to transfect cells. An equivalent amount of total protein was loaded in each soluble fraction lane and equivolume insoluble fractions were loaded. Blots were probed with antibodies specific for β-gal, 6xHis and GAPDH as indicated.

Ectopic expression of R4-uAb mediates proteasomal degradation. Having reduced the overall expression of β-gal in HEK293T cells such that it was no longer being partitioned to the insoluble fraction, R4-uAb-mediated silencing was re-evaluated. HEK293T cells were transiently transfected with pcDNA3-β-gal only (at 0.05 μg DNA) or co-transfected with pcDNA3-R4-uAb (0.05-1.25 μg DNA). Cellular β-gal levels were measured by immunoblotting (**Fig. 2.13**) and by determining β-gal activity 24 h post-transfection (**Fig. 2.15**). When HEK293T cells were co-transfected with pcDNA3-β-gal and increasing amounts of pcDNA3-R4-uAb, the β-gal levels were systematically reduced to as low as 3% of the steady-state levels measured in

cells transfected with only the pcDNA3-β-gal plasmid (**Fig. 2.13**). In contrast, no reduction in β-gal expression was observed following co-transfection with pcDNA3-D10-uAb or pcDNA3-scFv13-R4 plasmids (**Fig. 2.13**). Interestingly, cells co-transfected with R4-uAb^{R272A} again exhibited an intermediate level of β-gal expression (**Fig. 2.13**) as previously shown (**Fig. 2.8**). Importantly, the levels of a housekeeping protein, GAPDH, and a native binding partner of full-length CHIP, Hsp70, were not affected by co-transfections (**Fig. 2.13**). Additionally, evaluation of insoluble fractions confirmed that β-gal depletion was not due to cellular partitioning (**Fig. 2.13**).

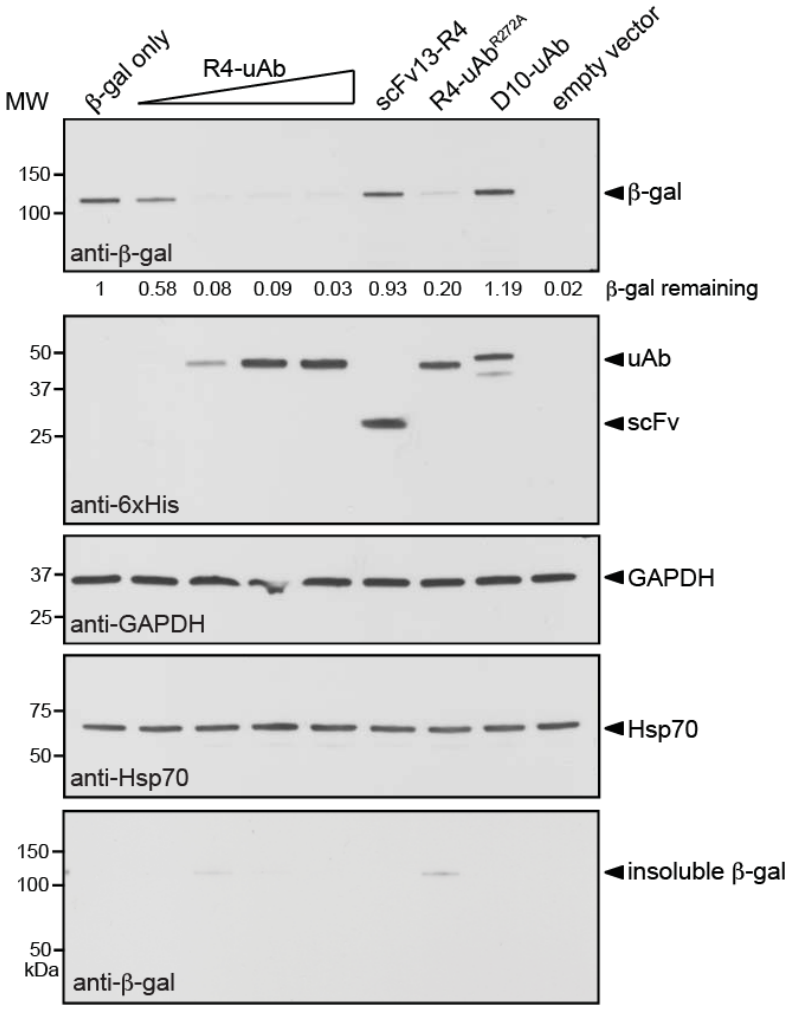


Figure 2.13 R4-uAb-mediated proteolysis of β-gal in HEK293T cells. Immunoblots of extracts prepared from HEK293T cells transfected with pcDNA3-β-gal alone (β-gal only) or co-transfected with one of the following: pcDNA3-R4-uAb, pcDNA3-D10-uAb, pcDNA3-R4-uAb^{R272A}, or pcDNA3-scFv13-R4. The triangle indicates increasing amounts of pcDNA3-R4-uAb plasmid DNA used to transfect cells. An equivalent amount of total protein was loaded in each lane. Blots were probed with antibodies specific for β-gal, 6xHis, GAPDH and Hsp70 as indicated. The immunoblot results are representative of at least three replicate experiments.

Similar results for co-transfection of pcDNA3-β-gal with pcDNA3-R4-uAb were obtained in BHK21 and COS-7 cells (**Fig. 2.14**), indicating that targeted protein silencing by engineered ubiquibodies is transferable between different cell lines. Notably, in both BHK21 and COS-7 cell lines, the one-to-one ratio of co-transfection with pcDNA3-β-gal and pcDNA3-R4-uAb (lowest dosage) was as effective at silencing β-gal as higher ratios of ubiquibody. Additionally,

higher amounts of pcDNA3-R4-uAb appeared to partition some β-gal to the insoluble fractions (**Fig. 2.14** insoluble) even at the reduced pcDNA3-β-gal transfection level. Thus the underlying physiological explanation for sub-cellular partitioning may still be significant.

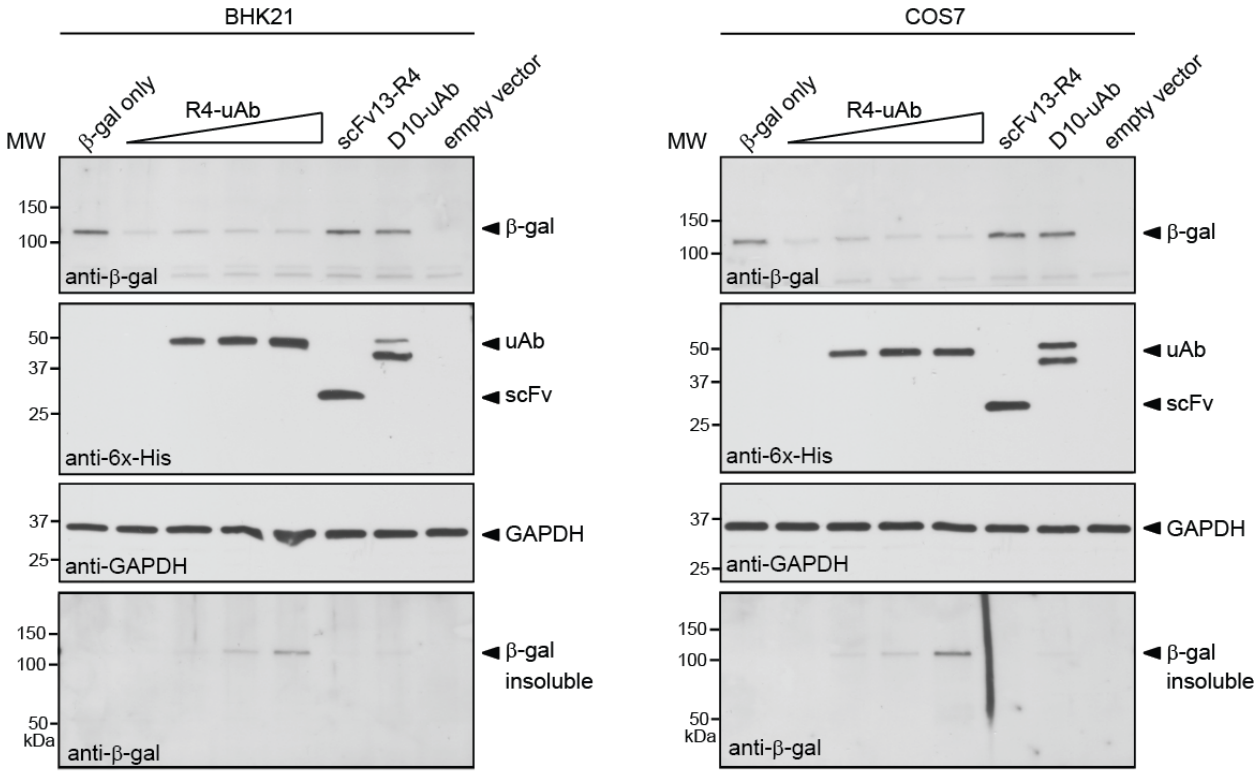


Figure 2.14 R4-uAb-mediated proteolysis of β-gal in different cell lines. (a) Immunoblots of extracts prepared from BHK21 (left) and COS-7 (right) cells transfected with pcDNA3-β-gal alone (β-gal only) or co-transfected with pcDNA3-β-gal along with one of the following: pcDNA3-R4-uAb, pcDNA3-scFv13-R4, or pcDNA3-D10-uAb. The triangle indicates increasing amounts of pcDNA3-R4-uAb plasmid DNA used to transfect cells. An equivalent amount of total protein was loaded in each lane. Blots were probed with antibodies specific for β-gal, 6x-His and GAPDH as indicated. The immunoblot results are representative of two replicate experiments.

In all three cell lines, β-gal knockdown was also detectable using β-gal activity assays (**Fig. 2.15**). Biological triplicate co-transfections using the best knockdown levels in each cell line (i.e. highest level of pcDNA3-R4-uAb in HEK293T and lowest level of pcDNA3-R4-uAb in BHK21 and COS-7) showed reproducible, significant knockdown compared to transfection with

pcDNA3-β-gal alone and control co-transfections with either pcDNA-scFv13-R4 or pcDNA3-D10-uAb (**Fig. 2.15**).

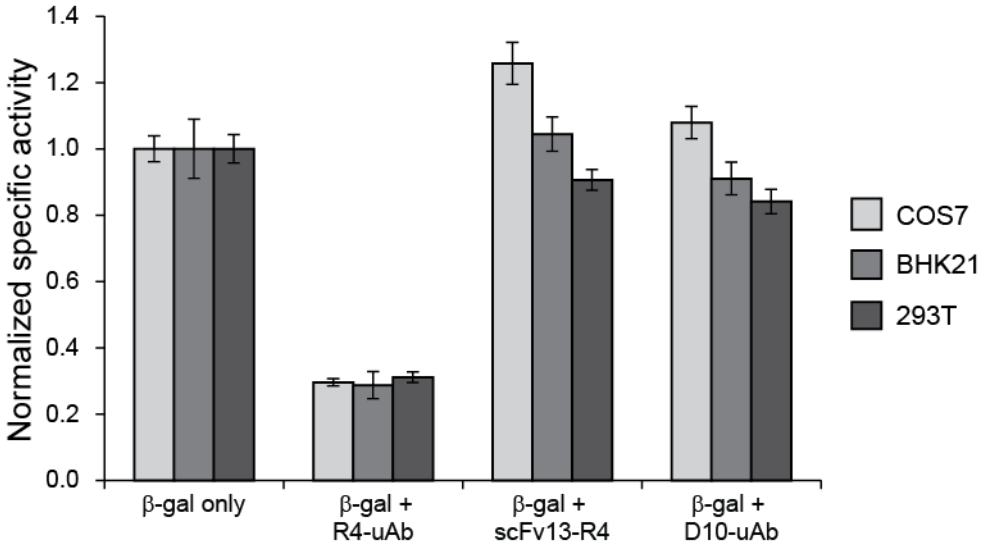


Figure 2.15 β-gal activity assay detecting knockdown in multiple cell lines. β-gal activity measured in samples transfected as described in Figures 2.13 and 2.14. In HEK293T cells the highest level of co-transfection was tested, whereas in COS-7 and BHK21, the lowest co-transfection level was used (control co-transfections were performed with the same amount of DNA). Data was normalized to the signal for the β-gal only control and is expressed as the mean ± standard deviation of the mean (SDM) of biological triplicates including error propagation from the activity assay absorbance and the total protein assay used in calculations.

Finally, to confirm that R4-uAb specifically binds and ubiquitinates β-gal *in vivo*, β-gal interactions were evaluated using a pull-down assay. HEK293T cells were transiently transfected with pcDNA3-R4-uAb or pcDNA3-β-gal alone, or co-transfected with pcDNA3-β-gal and pcDNA3-R4-uAb, pcDNA3-scFv13-R4 or pcDNA3-D10-uAb. Each ubiquibody or scFv control contains a C-terminal 6xHis tag which was then used to isolate these proteins using Ni^{2+} affinity magnetic agarose beads. As expected, both R4-uAb and scFv13-R4, but not D10-uAb, co-precipitated β-gal (**Fig. 2.16**). Furthermore, high molecular weight proteins co-precipitated by R4-uAb were observed to cross-react with an anti-ubiquitin antibody (**Fig. 2.16**), suggesting that ubiquitinated β-gal was present in cells co-transfected with pcDNA3-R4-uAb. These results

suggest that β-gal depletion in mammalian cells results from specific target binding and ubiquitination by the engineered R4-uAb.

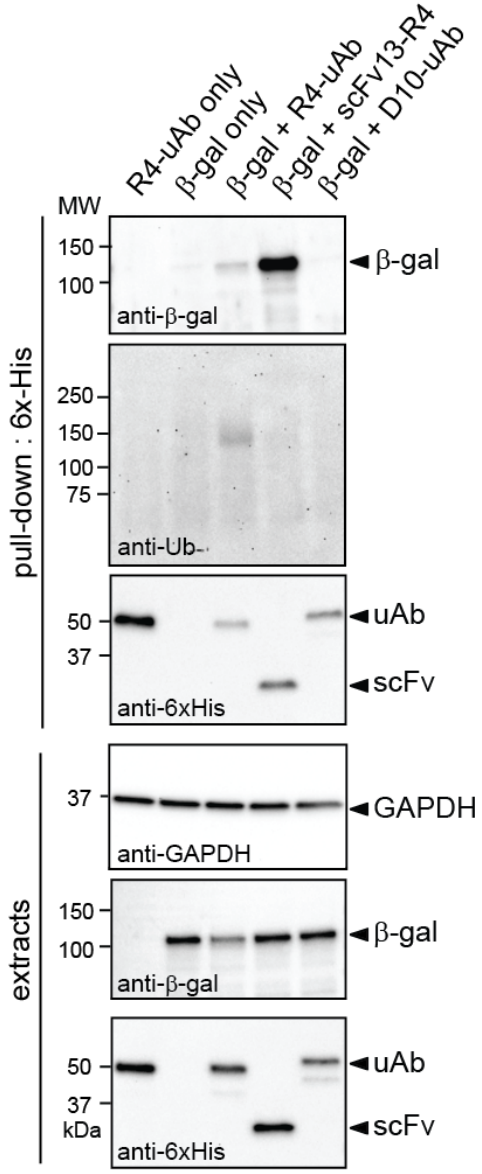


Figure 2.16 Co-precipitation of ubiquitinated β-gal in HEK293T cells. Immunoblots of pull-down samples or extracts prepared from HEK293T cells transfected with pcDNA3-R4-uAb alone (R4-uAb only), pcDNA3-β-gal alone (β-gal only) or co-transfected with pcDNA3-β-gal along with one of the following: pcDNA3-scFv13-R4, pcDNA3-R4-uAb, or pcDNA3-D10-uAb. Pull-down was performed by subjecting extracts to Ni-NTA magnetic agarose beads followed by immunoblotting with antibodies specific for β-gal, ubiquitin and 6x-His as indicated. An equivalent amount of total protein was loaded in each lane, as confirmed by immunoblotting with anti-GAPDH antibodies. Immunoblots are representative of three replicate experiments.

Discussion

Thus far, we have redirected the E3 ubiquitin ligase CHIP to target an otherwise stable protein, β-galactosidase, for degradation by the ubiquitin proteasome pathway. In creating this ubiquibody, we have exploited the modular architecture and intrinsic flexibility of CHIP [48]. Owing to the catalytic nature of ubiquitination, a single ubiquibody molecule has the potential for elimination of numerous copies of its target protein. This represents a major advantage over target inactivation using binding proteins alone, which require a one-to-one stoichiometric ratio (or greater) with their target because there are no elimination pathways for binding protein-target complexes.

While sub-cellular partitioning of β-gal in the presence of the ubiquibody was unexpected, the biological explanation may be elegantly simple. CHIP has been shown to interact with the chaperone-assisted selective autophagy pathway (CASA) [57]. Furthermore, it is known that K63-linked polyUb chains are recognized by CASA receptors such as p62 which then target aggregated proteins for autophagy and lysosomal degradation [58]. Notably, the R272A point mutation which was used to inhibit E2 interaction may not inhibit CHIP's interactions with the Uev1a/Ubc13 heterodimer E2 which forms K63-linked chains in concert with wild-type CHIP [59]. Thus while further experimentation will be required to prove whether the insoluble partitioning was due to autophagy pathway processing, it could be a valuable alternate approach for targeted degradation of oligomeric (such as β-gal which forms a tetramer) or aggregation prone substrates. To further investigate this phenomenon, immunofluorescent labeling could be used to label co-transfected cells for both β-gal and standard autophagy or lysosomal markers (i.e. LC3B or cathepsin L respectively).

In this work, we have attained significant knockdown in three different cell lines (HEK293T, BHK21 and COS-7) using transient co-transfection to deliver target and ubiquibody proteins. However, it is anticipated that infection with different doses of recombinant adenoviruses expressing ubiquibody genes could lead to complete ablation of target protein levels by overcoming the lower transfection efficiency that is typically associated with transient co-transfection [23]. Furthermore, we have so far only addressed the ubiquitination of one novel substrate, namely β-galactosidase. Thus the next step towards developing a generalizable protein knockdown technology is to swap the scFv13-R4 domain with other designer binding proteins (DBPs) and evaluate ubiquibody functionality.

Materials and Methods

Plasmid construction. Full-length human CHIP (a gift from Cam Patterson) was PCR amplified for cloning into pET28a(+) using 5' *Nco*I and 3' *Sal*I restriction sites. DNA encoding a double tag of FLAG-6xHis was created by dimerizing primers with a 5' *Sal*I overhang and a 3' *Hin*dIII overhang. Double ligation was performed to insert the CHIP PCR product and phosphorylated primer dimer between *Nco*I and *Hin*dIII sites in pET28a(+), yielding plasmid pET28a-CHIP. To create truncated CHIPΔTPR, DNA corresponding to amino acids 128-303 of human CHIP was PCR amplified with introduction of a 5' *Nco*I site and a 3' *Sal*I site. Double ligation was performed as above to insert this product along with the primer dimer into pET28a(+), yielding pET28a-CHIPΔTPR. The genes encoding scFv13 and scFv13-R4 (a gift from Pierre Martineau) were PCR amplified and again each was double ligated into pET28a(+) with the above primer dimer, yielding the control plasmids pET28a-scFv13 and pET28a-scFv13-R4, respectively. To create CHIPΔTPR fusions, PCR was used to introduce an *Eco*RI site followed by a short, flexible linker of GSGSG to the 5' end of CHIPΔTPR. In parallel, each of the DBPs including

scFv13, scFv13-R4, and the scFv D10 (a gift from Andreas Plückthun) was PCR amplified with a 5' *Nco*I site and 3' *Eco*RI site. Double ligation was then used to insert each scFv along with the GSGSG linker-CHIPΔTPR product between the *Nco*I and *Sal*I sites of pET28a-CHIP, yielding pET28a-scFv13-uAb, pET28a-R4-uAb, and pET28a-D10-uAb (**Fig. 2.2a**). The R272A point mutation was introduced into pET28a-R4-uAb using a QuikChange site-directed mutagenesis kit (Stratagene). For expression in eukaryotic cells, the above constructs were PCR amplified from their pET28a(+) backbones using primers that introduced a Kozak sequence at the start codon as well as a 5' *Hin*dIII site and a 3' *Xba*I site. The resulting PCR products were then cloned between the *Hin*dIII and *Xba*I sites of plasmid pcDNA3. The target substrate protein β-gal was PCR amplified using primers that introduced a Kozak sequence at the start codon as well as 5' *Xho*I and a 3' *Xba*I site and cloned in the corresponding sites of pcDNA3. For microscopy studies, the β-gal-eGFP fusion was created by PCR amplifying β-gal with a 5' *Xho*I site and a 3' *Bam*HI site which was then ligated into pcDNA3.1(+)-α-synuclein-eGFP (a gift from Anne Messer) replacing α-synuclein to create pcDNA3.1(+)-β-gal-eGFP.

Protein expression and purification. All purified proteins were obtained from cultures of *E. coli* BL21(DE3) cells grown in 500 mL of Luria-Bertani (LB) medium. Expression was induced with 0.1 mM IPTG when the culture density (Abs_{600}) reached 0.6-0.8 and proceeded at 30°C for 6 h, after which cells were harvested by centrifugation at 4,000xg for 20 min at 4°C. The resulting pellets were stored at -80°C overnight. Thawed pellets were resuspended in 15 mL buffer A (20 mM sodium phosphate, 0.5 M NaCl and 20 mM imidazole, pH 7.4) and lysed with a high-pressure homogenizer (Avestin EmulsiFlex C5). Lysates were cleared by centrifugation at 20,000xg for 20 min at 4°C and then subjected to Ni^{2+}-affinity purification with an ÅKTA FPLC using a 1-ml HisTrap column (GE Healthcare). Samples were washed with 10% buffer B

(20 mM sodium phosphate, 0.5 M NaCl and 500 mM imidazole, pH 7.4) before elution with 50% buffer B. Purified proteins were desalted over a 5-mL HiTrap column equilibrated with ubiquitination reaction buffer (20 mM MOPS, 100 mM KCl, 1 mM DTT, 5 mM $MgCl_2$, pH 7.2).

Enzyme-linked immunosorbent assay (ELISA). For ELISA, a previously established protocol was used to detect binding to β-gal [33], with slight modification. Briefly, a 96-well EIA plate was coated with 100 μL β-gal (Sigma) at 10 μg/mL overnight at 4°C. The plate was then washed three times with 200 μL PBST (1x PBS + 0.1% Tween20) per well for 5 min with shaking and blocked with 250 μL PBS with 3% milk per well at room temp, slowly mixing for 3 h. Following three washes as above, purified protein samples were introduced in blocking buffer as serial dilutions with 60 μL per well and incubated at room temp slowly mixing for 1 h. Three washes were used to remove non-bound protein before introducing 50 μL of anti-6x-His-HRP (diluted 1:10,000 in PBST + 1% milk) and incubating at room temp with slow mixing for 1 h. Three final washes were performed before incubation with 200 μL OPD (Sigma Fast tablets) in the dark for 30 min. The reaction was then quenched with 50 μL 3N H_2SO_4 and absorbance read at 492 nm.

***In vitro* ubiquitination assay.** Ubiquitination assays were performed as previously described [48] in the presence of 0.1 μM purified human recombinant UBE1 (Boston Biochem), 2 μM human recombinant UbcH5α/UBE2D1 (Boston Biochem), 3 μM uAb (or equivalent control protein), 3 μM *E. coli* β-gal (Sigma), 50 μM human recombinant ubiquitin (Boston Biochem), 4 mM ATP and 1 mM DTT in 20 mM MOPS, 100 mM KCl, 5 mM $MgCl_2$, pH 7.2. Reactions were carried out at 37°C for 2 h (unless otherwise noted) and stopped by boiling in 2x Laemmli loading buffer for analysis by immunoblotting.

Mass spectrometry analysis. For LC-MS/MS sample preparation, ubiquitination assays were performed as previously described [54] but with 0.1 μM UBE1, 20 μM UbcH5α, 20 μM R4-

uAb, 20 μM β-gal and 500 μM ubiquitin. Reactions were resolved by SDS-PAGE and stained by Coomassie prior to gel excision. The protein bands were cut from an SDS-PAGE gel and cut into ~1 mm cubes. The gel bands were washed in 200 μL DI water for 5 min, followed by 200 μL 100 mM ammonium bicarbonate (ambic)/acetonitrile (ACN) (1:1) for 10 min and finally 200 μL ACN for 5 min. The acetonitrile was discarded and the gel bands were dried in a speed-vac for 10 min. The gel pieces were rehydrated with 70 μL of 10 mM DTT in 100 mM ambic and incubated for 1 h at 56°C. The samples were allowed to cool to room temperature, after which 100 μL of 55 mM iodoacetamide in 100 mM ambic was added and the samples were incubated at room temp in the dark for 60 min. Following incubation, the gel slices were again washed as described above. The gel slices were dried and rehydrated with 50 μL trypsin at 50 ng/μL in 45 mM ambic, 10% ACN on ice for 30 min. The gel pieces were covered with an additional 25 μL of 45 mM ambic, 10% ACN and incubated at 37°C for 19 h. The digested peptides were extracted twice with 70 μl of 50% ACN, 5% formic acid (FA) (vortexed 30 min, sonicated 10 min) and once with 70 μl of 90% ACN, 5% FA. Extracts from each sample were combined and lyophilized.

The lyophilized in-gel tryptic digest samples were reconstituted in 20 μL of nanopure water with 0.5% FA for nanoLC-ESI-MS/MS analysis, which was carried out by a LTQ-Orbitrap Velos mass spectrometer (Thermo-Fisher Scientific) equipped with a CorConneX nano ion source device (CorSolutions LLC). The Orbitrap was interfaced with a nano HPLC carried out by an UltiMate3000 UPLC system (Dionex). The gel extracted peptide samples (2-4 μL) were injected onto a PepMap C18 trap column-nano Viper (5 μm, 100 μm × 2cm, Thermo Dionex) at 20 μL/min flow rate for on-line desalting and then separated on a PepMap C18 RP nano column (3 μm, 75 μm x 15 cm, Thermo Dionex) which was installed in the “Plug and Play” device with

a 10-μm spray emitter (NewObjective). The peptides were then eluted with a 90 min gradient of 5% to 38% ACN in 0.1% FA at a flow rate of 300 nL/min. The Orbitrap Velos was operated in positive ion mode with nano spray voltage set at 1.5 kV and source temperature at 275°C. Internal calibration was performed with the background ion signal at m/z 445.120025 as the lock mass. The instrument was operated in parallel data-dependent acquisition (DDA) mode using FT mass analyzer for one survey MS scan for precursor ions followed by MS/MS scans on top 7 highest intensity peaks with multiple charged ions above a threshold ion count of 7500 in both LTQ mass analyzer and HCD-based FT mass analyzer at 7,500 resolution. Dynamic exclusion parameters were set at repeat count 1 with a 15 sec repeat duration, exclusion list size of 500, 30 sec exclusion duration, and ±10 ppm exclusion mass width. HCD parameters were set at the following values: isolation width 2.0 m/z, normalized collision energy 35%, activation Q at 0.25, and activation time 0.1 msec. All data were acquired using Xcalibur 2.1 operation software (Thermo-Fisher Scientific).

All MS and MS/MS raw spectra were processed and searched using Proteome Discoverer 1.3 (PD1.3, Thermo-Fisher Scientific) against databases downloaded from NCBI-nr database. The database search was performed with two-missed cleavage site by trypsin allowed. The peptide tolerance was set to 10 ppm and MS/MS tolerance was set to 0.8 Da for CID and 0.05 Da for HCD. A fixed carbamidomethyl modification of cysteine, variable modifications on methionine oxidation, and ubiquitin modification of lysine were set. The peptides with low confidence score (with Xcorr score <2 for doubly charged ion and <2.7 for triply-charged ion) defined by PD1.3 were filtered out and the remaining peptides were considered for the peptide identification with possible ubiquitination determinations. All MS/MS spectra for possibly identified ubiquitination peptides from initial database searching were manually inspected and

validated using both PD1.3 and Xcalibur 2.1 software.

Cell culture, transfection and lysate preparation. All cell lines were obtained from ATCC and cultured in standard medium at 37°C with 5% CO_2. HEK293T and COS-7 cells were cultured in DMEM with 10% heat inactivated FBS and 1% antibiotic-antimycotic (Cellgro). BHK21 cells were cultured in EMEM with 10% heat inactivated FBS and 1% antibiotic-antimycotic (Cellgro). Cells were transfected in 6-well plates at 60-80% confluency with jetPRIME® (Polyplus Transfection) and at 4 h post-transfection the growth media was refreshed. For HEK293T cells, a 1:2 DNA to jETPRIME® (w:v) ratio was used for transfection and 2 μg total DNA was transfected per well. For BHK21 cells, a 1:3 DNA to jETPRIME® (w:v) ratio was used for transfection and 2 μg total DNA was transfected per well. And for COS-7 cells, a 1:3 DNA to jETPRIME® (w:v) ratio was used for transfection and 1 μg total DNA was transfected per well. In all experiments, empty pcDNA3 plasmid was used to balance transfection levels across samples to reach the total DNA level as specified per cell line.

At 24 h post-transfection, cells were harvested by trypsinization, washed with PBS and frozen at -20°C until analyzed by immunoblotting. Thawed cells were lysed in NP40 lysis buffer (150 mM NaCl, 1% Nonidet P-40 and 50 mM TrisHCl, pH 7.4) by pipetting and mixing at 4°C for 30 min, followed by removal of the insoluble fraction at 18,000xg at 4°C for 20 min. Insoluble pellets were then washed with 50mM TrisHCl and 1mM EDTA, pH 8 followed by solubilization in an equal volume of 2% SDS in PBS by boiling for 10 min. Cooled samples were centrifuged at room temp for 10 min at 13,200 RPM to remove remaining cellular debris. Both soluble and insoluble fractions were boiled in 2x Laemmli sample buffer for analysis by immunoblotting. Soluble fraction lysates were normalized using a detergent compatible total protein assay (Bio-Rad) and 10 μg total protein was loaded with equivolume insoluble fractions

for immunoblotting comparison.

Protein analysis. SDS-PAGE and immunoblotting of proteins was performed according to standard procedures. BioRad Coomassie G-250 stain was used to visualize proteins in SDS-PAGE (BioRad, Mini-PROTEAN® TGX). The following primary antibodies were utilized for immunoblotting: rabbit anti-β-gal (Abcam, ab616), mouse anti-ubiquitin (Millipore, P4D1-A11), rabbit anti-Lys48 (Millipore, Apu2), rabbit anti-6x-His-HRP (Abcam, ab1187), mouse anti-GAPDH (Millipore, 6C5), mouse anti-FLAG®-HRP (Sigma, M2), mouse anti-Hsp70 (Enzo Life Sciences, C92F3A). Secondary antibodies goat anti-rabbit IgG (H+L) and anti-mouse IgG (H+L) with HRP conjugation (Promega) were utilized as needed.

Pull-down assays. Thawed cells were lysed as above; then clarified lysates were normalized by a detergent compatible total protein assay (Bio-Rad) and diluted to contain 0.75 mg/mL of protein with 10 mM imidazole to reduce non-specific binding. 200 μL of diluted lysates were incubated with 30 μL of Ni-NTA magnetic agarose beads (Qiagen, 5% solution) mixing at 4°C for 1-2 h and washed with 20 mM imidazole in lysis buffer. Bound proteins were eluted with 250 mM imidazole (50 mM TrisHCl pH 7.9 and 50 mM NaCl) and boiled in 2x Laemmli sample buffer for analysis by immunoblotting.

β-gal activity assay. β-gal activity was determined using a β-gal assay kit (Invitrogen) according to the manufacturer's instructions for the microtiter plate format. Briefly, cell pellets were resuspended in 1x lysis buffer (0.25 M Tris, pH 8.0) and lysed with three freeze-thaw cycles before clarification at 18,000xg for 5 min at 4°C. Then 2.5 μL HEK293T lysate or 10 mL COS7 or BHK21 lysate was added to a well containing 50 μL 1x cleavage buffer (0.1 M sodium phosphate, 10 μM KCl, 1 μM $MgSO_4$-$7H_2O$, pH 7) with β-mercaptoethanol and 17 μL ONPG (4 mg/mL). Reactions proceeded at 37°C for 30 min and were stopped with 125 μL stop buffer (1

M sodium carbonate) before measuring absorbance at 420 nm. The amount of ONPG hydrolyzed was calculated using the following formula: nmoles ONPG hydrolyzed = $(Abs_{420})(1.92x10^5$ nl) / (4500 nl/nmole-cm)(1 cm). Specific activity was determined according to the following formula: specific activity = ONPG hydrolyzed/t/mg protein; where t is the reaction time in minutes and mg is the amount of total protein assayed. The background activity from untransfected cell lysates was subtracted for each sample and specific activities within biological samples were normalized to cells transfected with β-gal alone.

Microscopy and Immunofluorescent labeling. Live cell GFP fluorescence imaging was performed with a Zeiss Axiovert A1 with LD Plan-Neofluar 40x (0.6 numerical aperture) objective. For immunofluorescent studies, HEK293T cells were grown on glass coverslips coated with poly-L-lysine, fixed with 3.7% formaldehyde, permeabilized by 0.2% Triton X-100 and blocked in 3% fetal bovine serum. Primary anti-β-gal antibody (Pierce) incubation was done at room temperature for 1 h in a humidity chamber, followed by a 1 h incubation at room temperature with Alexa Fluor-labeled secondary antibody (AlexaFluor568 goat anti-rabbit IgG, Molecular Probes) at room temperature in the dark. Coverslips were mounted on VectaShield mounting media (H-1000) and imaged with the Zeiss Axiovert A1 and 10x EC Plan-Neofluar (0.3 numerical aperture) objective with an AxioCam ICm1 digital camera.

Acknowledgements

We thank Cam Patterson, Pierre Martineau, Anne Messer and Andreas Plückthun for kindly providing plasmids encoding genes used in this study. We thank Wei Chen and Sheng Zhang at the Cornell Biotechnology Resource Center for performing the mass spectrometry experiments and database searches. We also thank Pengbo Zhou and Jianxuan Zhang at Weill Cornell Medical College for initial training using HEK293T cells and transient transfection

methods. We thank Shu-Bing Qian in Nutritional Sciences for initial CHIP purification methods. And we thank Erin Stephens for providing technical aid and methodology for immunofluorescent labeling. This material is based upon work supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Institutes of Health Chemical Biology Training Grant T32 GM008500 (fellowship to A.D.P.) and the National Science Foundation GK-12 DGE-0841291 (fellowship to A.D.P.).

CHAPTER 3

EXPANDING SUBSTRATE SPECIFICITY WITH DESIGNER BINDING PROTEINS

Introduction

While antibodies and their fragments have been the workhorse for protein-protein interactions in biotechnology applications for years, there are many recent developments toward using more of nature's diverse binding domains for engineering applications. Two major classes of engineered binding proteins include immunoglobulin-like domains (e.g. monobodies and lipocalins) and domain repeat scaffolds (e.g. ankyrin and leucine-rich repeats). For example, monobodies are based on the human 10^{th} fibronectin type 3 domain (^{10}Fn3) which is structurally similar to the β-sandwich structure of the V_H domain of an antibody. As such, these domains contain three loops, which have been diversified to create libraries [44]. Notably, the ^{10}Fn3 domain does not require disulfide bonds and is devoid of natural free cysteines, making monobodies candidates for both intracellular expression and site specific modification using cysteine chemistry. The most developed type of repeat scaffold proteins are the designed ankyrin repeat proteins (DARPins). These utilize varying numbers of internal ankyrin repeats (AR), each with six randomized residues, and capped with consensus N- and C-terminal ARs which provide stability [39, 40].

In order to confirm the versatility, modularity and robustness of ubiquibodies, we further diversified the substrate binding domains fused to CHIPΔTPR by using various designer binding proteins (DBPs). The major design constraint in selecting DBPs for ubiquibody creation is that they must be functional in the reducing environment of the cytoplasm. Thus to test the modularity of our ubiquibody design, we have utilized intrabodies, which are selected for intracellular stability, and monobodies and DARPins (designed ankyrin repeat proteins), which

naturally do not contain disulfide bonds.

Results

Diversification of ubiquibodies with DBPs. Using the modular platform developed for ubiquibody expression in *E. coli* (see **Fig. 2.2a**), we first cloned a collection of scFvs fused to CHIPΔTPR to evaluate intracellular expression. As was predicted, only scFvs selected for intracellular stability were well expressed as ubiquibodies in the cytoplasm of *E. coli* (**Fig. 3.1a**). Notably, scFvs 3DX and HAG which were isolated by phage display and *E. coli* periplasmic expression respectively did not show soluble expression as fusions to CHIPΔTPR [60, 61]. Yet most intrabodies, including scFv13-R4, D10, and J21 created stable, soluble ubiquibodies when fused to CHIPΔTPR (**Fig. 3.1a**). Both scFv D10 and J21 were selected from a synthetic human combinatorial antibody library (HuCAL®) using a protein fragment complementation assay (PCA) in the cytoplasm of *E. coli* [51]. Clone D10 binds the bacteriophage capsid protein D (gpD), and J21 binds c-Jun N-terminal kinase 2 (JNK2) [51]. While the scFv GCN4 was selected from an intrabody library utilizing yeast two-hybrid display, this ubiquibody produced conflicting results of both soluble and insoluble expression (data not shown) [35].

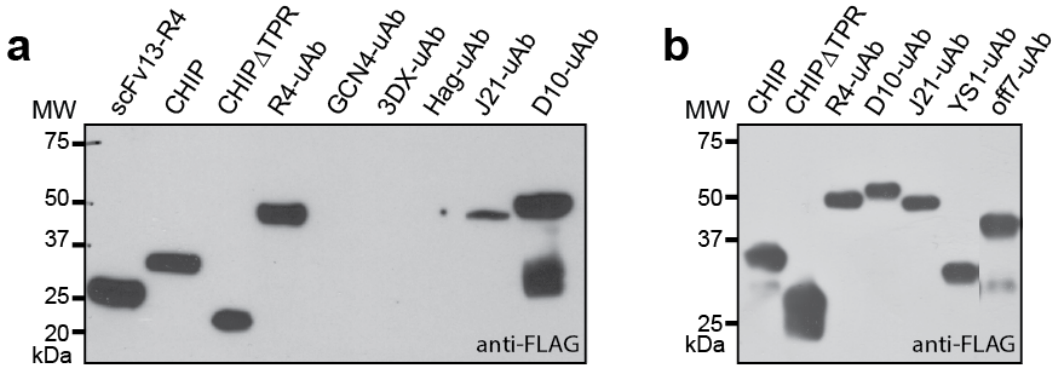


Figure 3.1 Soluble expression of diverse uAbs in *E. coli.* Western blot analysis of cell lysates isolated from *E. coli* strain BL21(DE3) expressing full-length scFv13-R4, CHIP, CHIPΔTPR, and various uAb constructs. (a) The scFv-uAbs were comprised of scFvs specific for *E. coli* β-gal (scFv13R4-uAb), *S. cerevisiae* GCN4 (GCN4-uAb), c-myc epitope (3DX-uAb), hemagglutinin epitope (Hag-uAb), JNK2 (J21-uAb) and bacteriophage gpD (D10-uAb) respectively. (b) Additional DBP-uAbs were comprised of the monobody YS1 and the DARPin off7, both specific for *E. coli* MBP (YS1-uAb and off7-uAb accordingly). An equivalent amount of total protein was loaded in each lane and anti-FLAG® antibodies were used to detect the expressed proteins.

Next we created ubiquibodies from non-immunoglobulin domains and compared their soluble expression in *E. coli* to intrabody based ubiquibodies (**Fig. 3.1b**). Specifically, we tested the monobody, YS1 and the DARPin, off7, both of which bind the *E. coli* maltose binding protein (MBP) and found that these smaller DBPs created ubiquibodies as soluble as wild-type CHIP (**Fig. 3.1b**). Interestingly, YS1, was isolated from a binary library built of tyrosine (Tyr, Y) and serine (Ser, S) residues in the binding domain loops, with the aim of studying conformational diversity rather than chemical diversity in protein-protein interactions [62, 63]. Alternatively, the DARPin utilized in this work, off7, was isolated from a combinatorial library of consensus-designed ankyrin repeat proteins of varying sizes, using *in vitro* ribosome display [40].

Functional testing of scFv-uAbs *in vitro*. We next tested the functionality of the most soluble scFv-uAbs, D10-uAb and J21-uAb, by purifying the ubiquibodies and their antigens from the *E. coli* strain BL21(DE3). The N-terminus of both antigens (gpD and JNK2) were augmented with

a double epitope tag of 6xHis and HA (hemagglutinin) to enable facile purification and differentiation from the uAbs by immunoblotting respectively. Notably neither of these tags contains lysine residues. Then we tested the binding affinity of D10-uAb to gpD and J21-uAb to JNK2 by ELISA (**Fig. 3.2**). The previously characterized R4-uAb, CHIPΔTPR or 5x1-uAb (a solubility improved scFv GCN4 uAb) were used as negative controls as they should not interact with either antigen. Interestingly, when the scFv D10 was fused to CHIPΔTPR the binding affinity for gpD improved slightly (**Fig. 3.2a**). CHIPΔTPR alone showed an affinity for gpD that was higher than expected, but which was identified to be due to the use of bovine serum albumin (BSA) as the blocking agent in the ELISA (compare to CHIPΔTPR signal in **Fig. 2.3a**, Abs_{492} < 0.2). The J21-uAb, however, did not maintain its minimal affinity for JNK2 (**Fig. 3.2b**). These results corroborate independent testing in our laboratory which found J21 to be a weak binding scFv which only exhibited differentiable binding affinity from non-binding control scFvs under optimized ELISA conditions. Furthermore, the J21 intrabody was isolated after just one round of selection versus the D10 intrabody which was isolated after ten rounds of competitive selection [51]. Additionally, the K_D for D10 was found to be 30.5 μM as determined by surface plasmon resonance (SPR) while the affinity for J21 could not be determined [51].

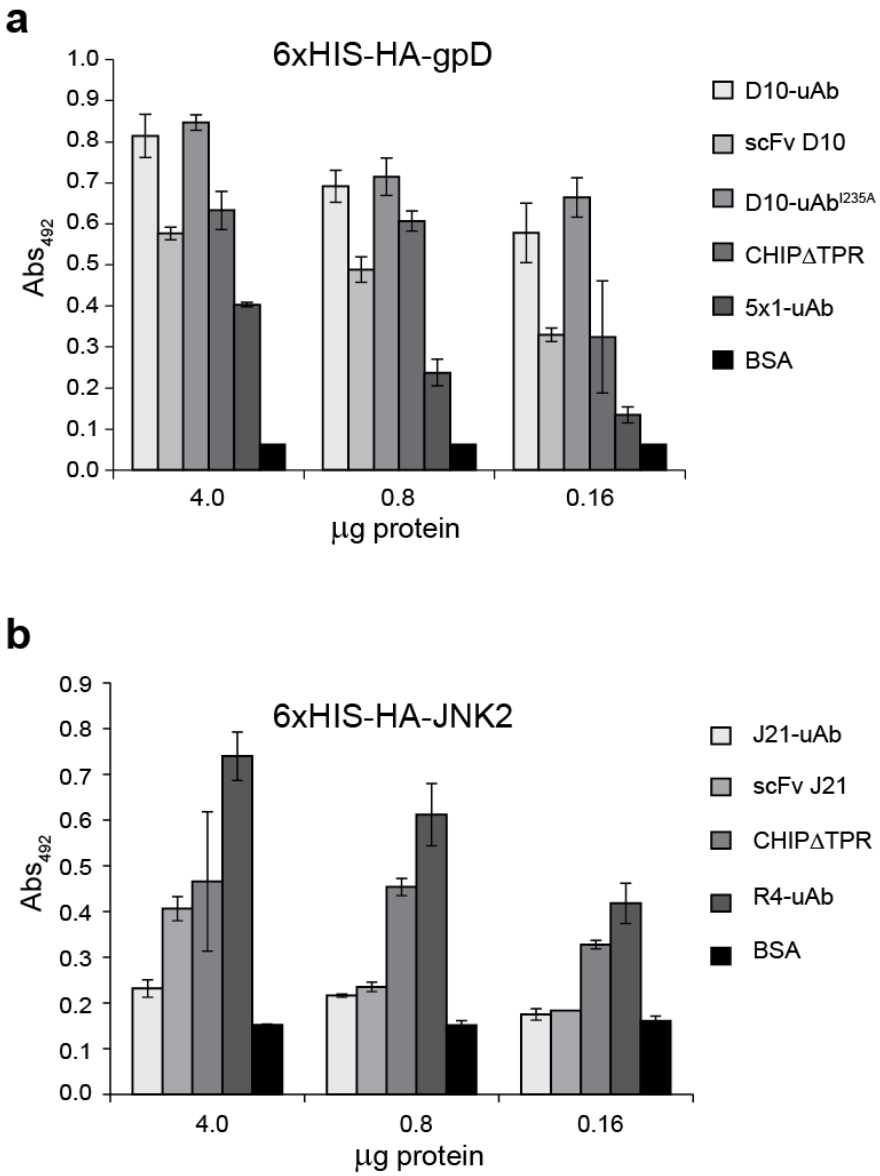


Figure 3.2 *In vitro* binding activity of scFv-uAbs as determined by ELISA. Binding activity of purified D10-uAb with gpD as antigen (a) and purified J21-uAb with JNK2 as antigen (b) as measured by ELISA. The D10 and J21 intrabodies served as positive controls while CHIPΔTPR and non-specific uAbs served as negative controls. Also tested was D10-uAb^{I235A}, a derivative of D10-uAb carrying a point mutation in the U-box domain that is known to block the interaction between CHIP and the E2 enzyme, UbcH5α. BSA which does not contain the epitope tag was used to show the background signal of the assay.

We next tested the scFv-uAbs for ubiquitin ligase activity by reconstituting the ubiquitination process *in vitro* using human UBE1 for the E1, UbcH5α as the E2 enzyme and 6xHis-HA-gpD, which has 6 surface lysines [64], or untagged JNK2, which has 20 surface

lysines [65], as the substrate proteins. Monoubiquitination of gpD was clearly evident when immunoblots were probed with anti-HA antibodies. Interestingly, the 6xHis-HA-gpD construct was not detected by the anti-6xHis antibodies despite analyzing ubiquitination reactions with increasing amounts of the substrate protein (**Fig. 3.3a**). The J21-uAb showed negligible ubiquitination of JNK2, despite using increased amounts of the ubiquibody (**Fig. 3.3b**). Notably, autoubiquitination of both scFv-uAbs was evident, consistent with wild-type CHIP autoubiquitination and which indicated that both ubiquibodies were active (**Fig. 3.3**) [48].

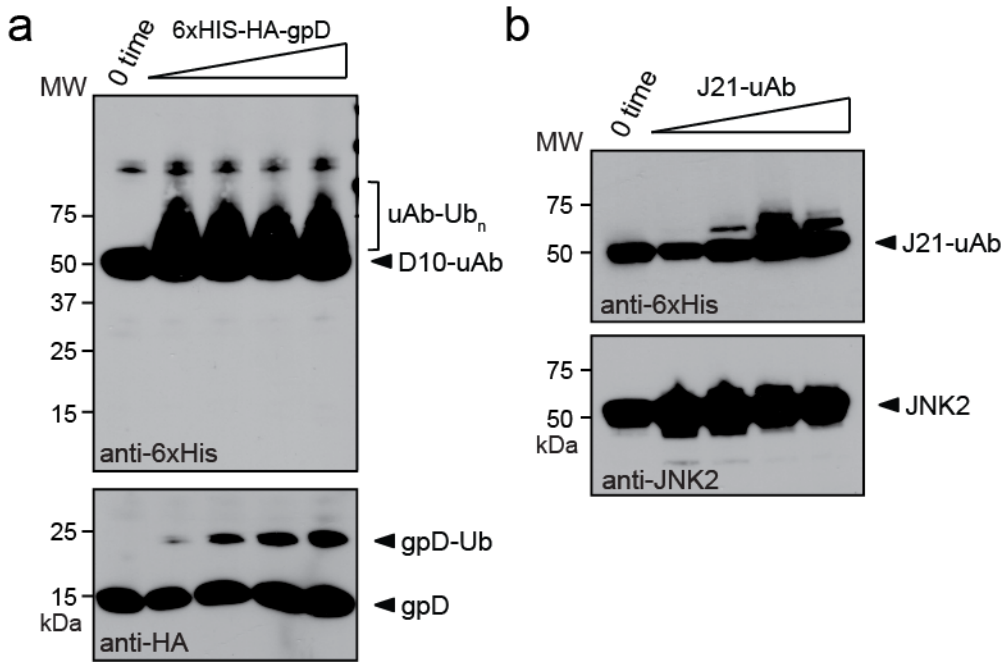


Figure 3.3 *In vitro* ubiquitination activity of scFv-uAbs. *In vitro* ubiquitination reactions with increasing amounts of the substrate protein gpD with the D10-uAb (a) or increasing amounts of the J21-uAb with JNK2 (b) were reacted for 2 h at 37°C. Reactions were stopped by boiling and immunoblotted with anti-6xHis, anti-HA and anti-JNK2 antibodies as indicated. Ubiquitination assay mixtures boiled prior to incubation were used as controls to identify unmodified proteins (0 time).

Further testing of D10-uAb with gpD revealed that polyubiquitination could be detected with antibodies specific for ubiquitin and K48-linked polyUb chains (**Fig. 3.4a**), confirming the presence of ubiquitin linkages that are known to signal proteasomal degradation [53]. When

reactions were performed with scFv D10 or CHIPΔTPR, there was no detectable ubiquitination of gpD. Likewise, D10-uAb^{I235A}, which carries a U-box domain point mutation to interfere with E2 binding, showed no modification of gpD, despite binding as well as the wild-type D10-uAb (**Fig. 3.2a** and **Fig 3.4a**). However, the identity of high molecular weight polyUb proteins was convoluted by the extensive autoubiquitination of both D10-uAb and CHIPΔTPR. Notably, neither D10-uAb nor CHIPΔTPR could be detected with anti-FLAG® antibodies (**Fig. 3.4a**), possibly due to ubiquitination of lysine residues within the epitope sequence. In evaluating the ubiquitination activity over time, it was evident that D10-uAb shows slower ubiquitination of gpD than wild-type CHIP does ubiquitinating Hsp70 (**Fig. 3.4b** and **c**). Finally, *in vitro* ubiquitination of gpD was monitored by SDS-PAGE and the formation of higher molecular weight species was used to confirm the presence of gpD-Ub$_n$ polyUb chains (**Fig. 3.4d**) which could not be detected directly with anti-HA nor anti-6xHis antibodies (**Fig. 3.3a**). Purified 6xHis-HA-gpD was compared to a complete ubiquitination reaction without incubation (0 time), the ubiquitination reaction after 2 h at 37°C, and a ubiquitination reaction lacking gpD, in order to account for E2 ubiquitination and uAb autoubiquitination (**Fig. 3.4d**). Minimally gpD modified with three ubiquitins could be detected, confirming polyubiquitination by D10-uAb.

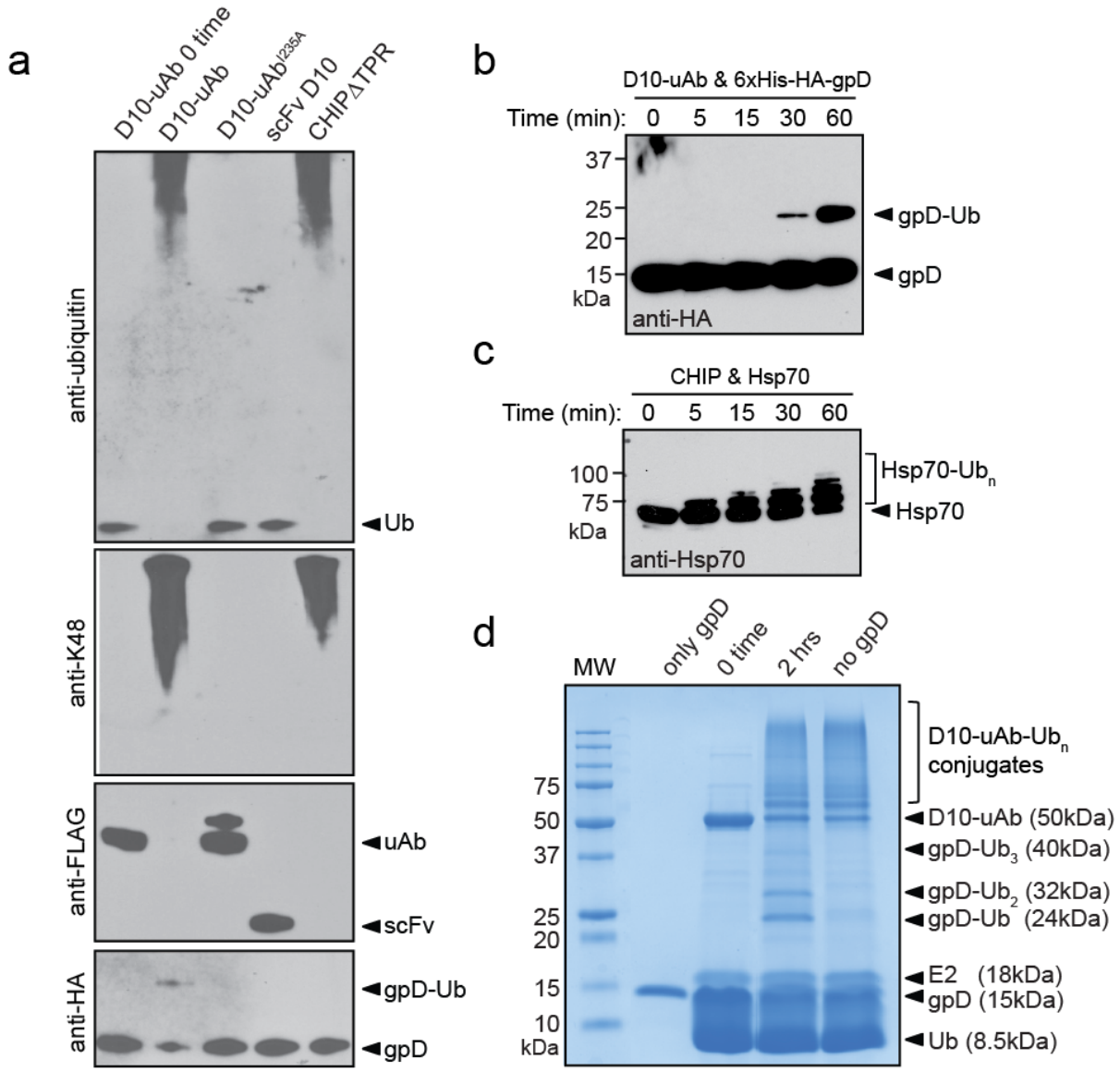


Figure 3.4 *In vitro* ubiquitination of gpD by D10-uAb. (a) Ubiquitination of gpD was carried out for 2 h with the controls D10-uAb^{I235A}, scFv D10, and CHIPΔTPR. An equivalent amount of total protein was added to each lane. Immunoblots were probed with anti-HA, anti-FLAG®, anti-K48 linked polyUb and anti-ubiquitin antibodies as indicated. (b & c) At the indicated times, *in vitro* ubiquitination reactions were stopped by boiling and immunoblotted with anti-HA and anti-Hsp70 antibodies. (d) Purified gpD (only gpD) and ubiquitination reactions at 0 h, 2 h and 2 h omitting gpD (no gpD) were analyzed by Coomassie staining. Protein bands corresponding to gpD, gpD-Ub conjugates, ubiquitin (Ub), E2, D10-uAb, D10-uAb-Ub conjugates and the molecular weight of the marker bands (MW) are indicated. An equivalent amount of total protein was loaded in each ubiquitination reaction lane. The purified gpD was loaded equivalent to the weight of gpD used in the ubiquitination reactions.

Ectopic co-expression of gpD with D10-uAb in mammalian cells. Having determined that D10-uAb could bind and ubiquitinate gpD *in vitro*, we next tested whether co-expression in mammalian cells would show targeted degradation of gpD. HEK293T cells were co-transfected with pCMV-HA-gpD, pcDNA3-D10-uAb and pGFP, a plasmid containing the GFP gene to evaluate transfection efficiency (**Fig. 3.5a**). Increasing amounts of D10-uAb appeared to reduce the levels of gpD expression while not affecting Hsp70 levels, a native binding partner of CHIP (**Fig. 3.5a**). Notably, there was no detectable level of D10-uAb expression (using anti-FLAG antibodies) without the co-expression of gpD (**Fig. 3.5a**). This was reproduced with a separate co-transfection (data not shown). Unfortunately, the apparent knockdown of gpD *in situ* was unable to be reproduced due to inconsistent expression of HA-gpD. Furthermore, we utilized a pull-down assay to test whether gpD was even interacting with D10-uAb *in situ* (**Fig. 3.5b**). HEK293T cells were transiently transfected with pCMV-HA-gpD alone, or co-transfected with pcDNA3-D10-uAb, pcDNA3-scFv D10 or pcDNA3-R4-uAb, which was used as a negative control. Each ubiquibody or scFv contained a C-terminal 6xHis tag which was then used to precipitate these proteins using Ni^{2+} affinity magnetic agarose beads. Unfortunately, neither the scFv D10 nor D10-uAb was able to co-precipitate HA-gpD *in situ*. These results draw attention to the affinity required of ubiquibodies for effective target binding and ubiquitination in the eukaryotic cellular context.

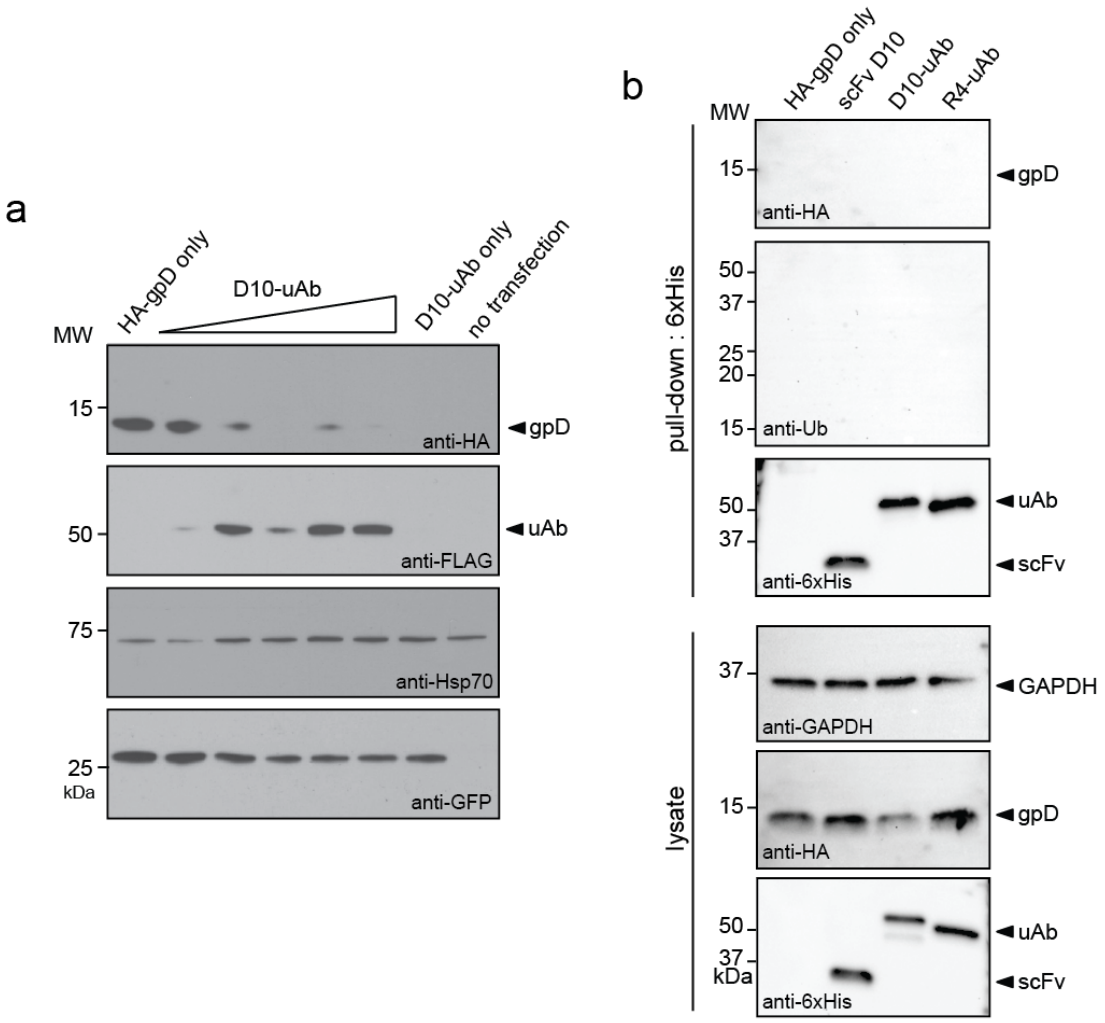


Figure 3.5 Ectopic co-expression of gpD and D10-uAb in mammalian cells. (a) Immunoblots of extracts prepared from HEK293T cells co-transfected with pGFP (transfection control) and pCMV-HA-gpD (HA-gpD only) or pcDNA3-D10-uAb (D10-uAb only), or co-transfected with all three plasmids. The triangle indicates increasing amounts of pcDNA3-D10-uAb plasmid DNA used to co-transfect cells. An equivalent amount of total protein was loaded in each lane and blots were probed with antibodies specific for HA, FLAG, Hsp70 and GFP as indicated. (b) Immunoblots of pull-down samples or extracts prepared from HEK293T cells transfected with pCMV-HA-gpD alone (gpD only) or co-transfected with one of the following: pcDNA3-D10, pcDNA3-D10-uAb or pcDNA3-R4-uAb. Pull-down was performed using Ni-NTA magnetic agarose beads followed by immunoblotting with antibodies specific for HA, ubiquitin and 6xHis as indicated. An equivalent amount of total protein was loaded in each lane, as confirmed by immunoblotting with anti-GAPDH antibodies.

Targeting maltose binding protein for degradation with DBP-uAbs. Moving away from the scFv-uAbs, we next tested our monobody and DARPin based ubiquibodies which were both

specific for the *E. coli* maltose binding protein (MBP). These uAbs were tested directly *in situ* by co-transfection of HEK293T cells with pcDNA3-MBP and pcDNA3-off7-uAb (**Fig. 3.6a**) or pcDNA3-YS1-uAb (**Fig. 3.6b**). While the DARPin based off7-uAb did not discernibly reduce the levels of MBP expression (**Fig 3.6a**), increasing amounts of the monobody based YS1-uAb systematically reduced MBP levels (**Fig.3.6b**). Furthermore, non-specific control ubiquibodies D10-uAb and R4-uAb did not affect MBP expression levels (**Fig. 3.6**).

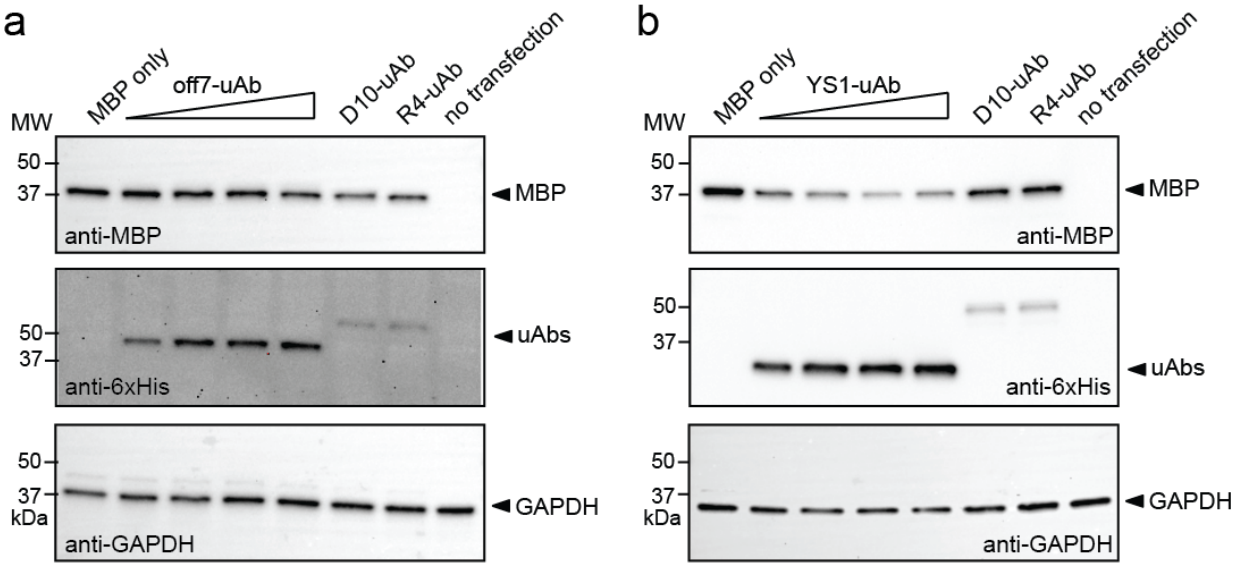


Figure 3.6 Comparing off7-uAb and YS1-uAb targeted proteolysis of MBP. Immunoblots of extracts prepared from HEK293T cells transfected with pcDNA3-MBP alone (MBP only) or co-transfected with pcDNA3-off7-uAb (a), pcDNA3-YS1-uAb (b), pcDNA3-D10-uAb, or pcDNA3-R4-uAb. The triangle indicates increasing amounts of plasmid DNA used to transfect cells. An equivalent amount of total protein was loaded in each lane, as confirmed with the loading control GAPDH. Blots were probed with antibodies specific for MBP, 6xHis, and GAPDH as indicated.

Confirming the YS1-uAb mediated silencing of MBP, replicate co-transfections with pcDNA3-MBP and increasing amounts of pcDNA3-YS1-uAb significantly reduced MBP levels compared to cells transfected with only the pcDNA3-MBP plasmid (**Fig. 3.7a)**. When HEK293T cells were co-transfected with control plasmids expressing YS1 or the non-specific R4-uAb, MBP depletion was not observed (**Fig. 3.7a** and **b**), confirming that YS1-uAb-mediated

proteolysis is CHIPΔTPR-dependent and ubiquibodies are highly specific. Additionally, Hsp70 levels were unchanged in the presence of YS1-uAb (**Fig. 3.7a**), demonstrating the remodeled specificity of CHIPΔTPR. Finally, *in situ* binding of MBP by YS1 and YS1-uAb was confirmed by co-precipitation (**Fig. 3.7b**); and the pull-down with YS1-uAb was enriched with ubiquitinated proteins. These results suggest that MBP depletion in mammalian cells results from specific target binding and ubiquitination by the engineered YS1-uAb.

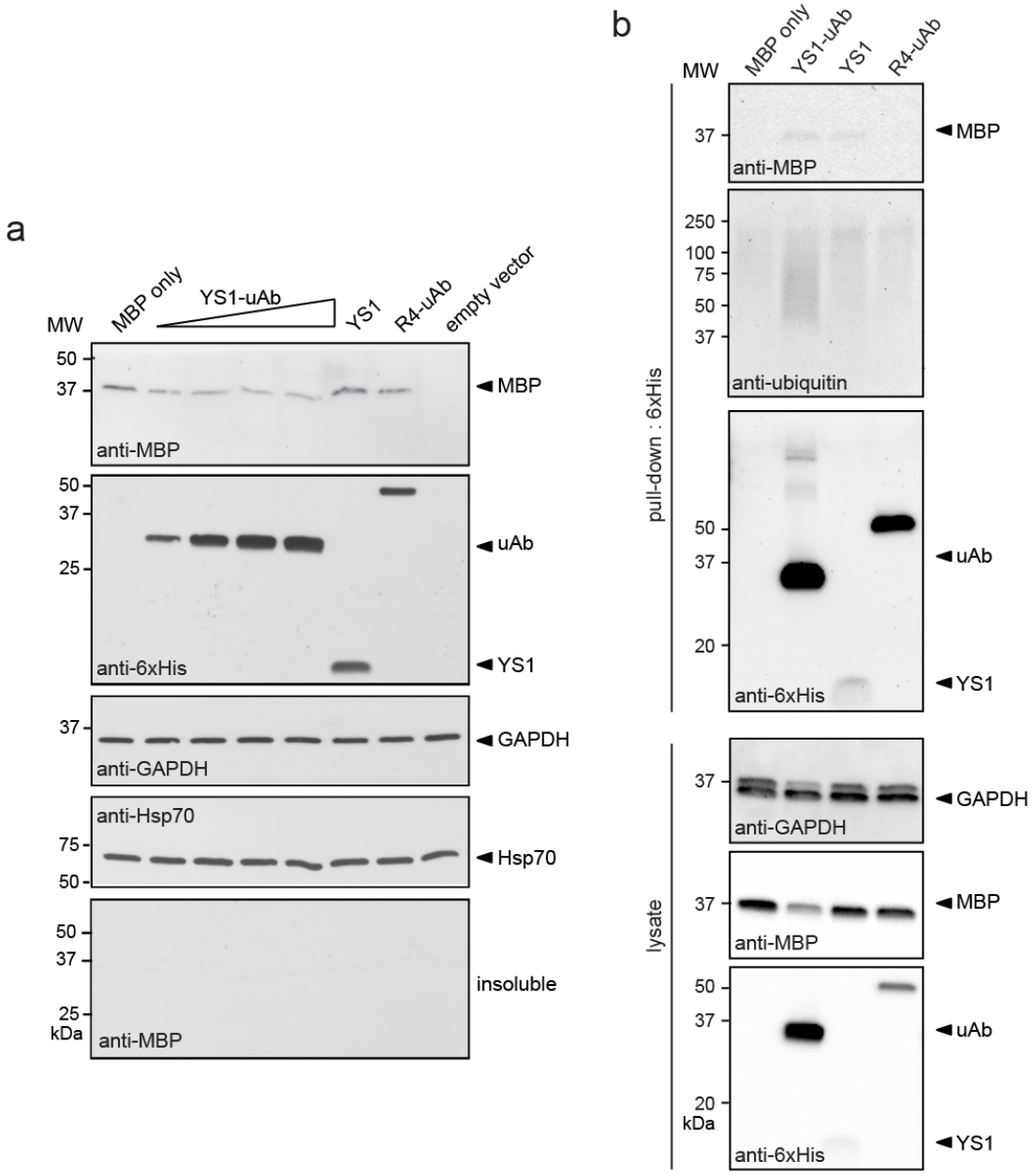


Figure 3.7 YS1-uAb-mediated proteolysis of MBP in mammalian cells. (a) Immunoblots of extracts prepared from HEK293T cells transfected with pcDNA3-MBP alone (MBP only) or co-transfected with pcDNA3-YS1-uAb, pcDNA3-YS1, or pcDNA3-R4-uAb. The triangle indicates increasing amounts of pcDNA3-YS1-uAb plasmid DNA used to transfect cells. An equivalent amount of total protein was loaded in each lane. Blots were probed with antibodies specific for MBP, 6xHis, Hsp70 and GAPDH as indicated. The immunoblot results are representative of at least three replicate experiments. (b) Immunoblots of pull-down samples or extracts prepared from HEK293T cells transfected with pcDNA3-MBP alone (MBP only) or co-transfected with pcDNA3-YS1, pcDNA3-YS1-uAb or pcDNA3-R4-uAb. Pull-down was performed using Ni-NTA magnetic agarose beads followed by immunoblotting with antibodies specific for MBP, ubiquitin and 6xHis as indicated. An equivalent amount of total protein was loaded in each lane, as confirmed by immunoblotting with anti-GAPDH antibodies.

Discussion

Domain swapping with different DBPs (e.g., intrabody and monobody) confirmed the conformational flexibility of CHIP in ubiquitin transfer [48], targeting multiple acceptor lysines on three structurally distinct substrates. These results also confirm the potential of CHIP for customizable target degradation in mammalian cells. Indeed, given the plethora of existing DBPs against known cellular targets and the availability of robust technologies for on-demand isolation of new DBPs that function inside cells [34, 51, 66], ubiquibodies are likely to become a powerful tool for reverse genetics.

Thus far, we considered the intracellular stability of a DBP as the major design constraint in creating ubiquibodies. However an additional design constraint may be the binding kinetics of DBPs. A comparison of the K_Ds of the DBPs and CHIP with its natural substrates Hsc70, Hsp70 and Hsp90 is given in **Table 3.1**. Notably, the two DBPs that showed consistent functionality in the eukaryotic cellular context, scFv13-R4 and YS1 have the closest K_Ds to those reported for CHIP. As well, over-expression of CHIP *in situ* shows faster degradation of Hsp70 than Hsc70, despite their high sequence similarity [67], most likely due to their sequence divergence at the C-terminus where CHIP binds these chaperones. This is also consistent with their hierarchal binding affinities with CHIP when compared under the same experimental set-up (ITC, [68]).

Table 3.1 Comparison of dissociation constants of CHIP and DBPs.

Binding Protein	**Substrate**	**K_D (μM)**	**k_{off} (s^{-1})/k_{on} (M^{-1} s^{-1})**	**Method, Source**
CHIP	Hsc70	0.07 (± 0.01)	$6x10^{-3}/8.7x10^{4}$	BLI [69]
CHIP	Hsc70	2.3 (± 0.3)	n.d.	ITC, [68]
CHIP	Hsp70	0.95 (± 0.01)	n.d.	ITC, [68]
CHIP	Hsp90	0.38 (± 0.04)	n.d.	ITC, [68]
scFv13-R4	β-gal	0.128 (± 0.049)	n.d.	Competitive ELISA, [66]
D10	gpD	30.5	n.d.	SPR, [51]
J21	JNK2	n.d. (50-100)	n.d.	Estimated, [51]
YS1	MBP	0.135	$7.6x10^{-2}/5.6x10^{5}$	SPR, [62]
off7	MBP	0.0044	$1.9x10^{-3}/4.2x10^{5}$	SPR, [40]

Also contributing to the significance of binding kinetics, E3 ubiquitin ligases must bind both a ubiquitin-charged-E2 and the substrate protein at the same time, creating a ternary complex necessary for substrate ubiquitination (**Fig. 3.8**) [70]. Then, the E3 must stay bound to the substrate long enough to enable poly-ubiquitin transfer or risk the substrate being deubiquitinated by deubiquitinating enzymes (DUBs) in the cellular milieu [26, 71]. This may put significant restraints on suitable k_{off} rates for ubiquibodies because the E2 charged with ubiquitin usually must cycle off and be replaced with another charged E2 in order for poly-ubiquitination to occur [72, 73]. Thus, it will be important that engineered ubiquitin ligases do not alter the E2 binding kinetics which defines processive ubiquitin chain synthesis ($\alpha K_{E2\sim Ub}$ and αK_{E2} in **Fig. 3.8**) [70, 73].

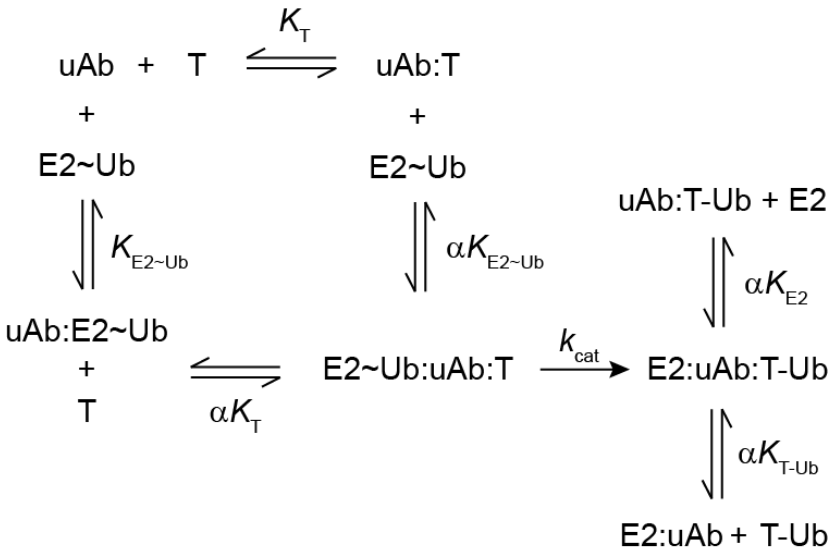


Figure 3.8 Enzyme scheme for ubiquitination kinetics. The kinetic model is adapted from Swinney et al. [70]. K_T and αK_T are dissociation constants of the ubiquibody for the target, T in the absence and presence, respectively, of E2 charged with ubiquitin, E2~Ub, in the U-box domain. $K_{E2\sim Ub}$ and $\alpha K_{E2\sim Ub}$ are dissociation constants of the ubiquibody for the ubiquitin charged E2 in the absence and presence, respectively, of the target in the DBP active site. A value of $\alpha = 1$ indicates that the equilibrium dissociation constants are not influenced by the binding of the alternative substrate. $K_{T\text{-}Ub}$ and K_{E2} are dissociation constants of the ubiquibody for the ubiquitinated target and E2 without ubiquitin, respectively [70].

Materials and Methods

Plasmid construction. Detailed methodology for the construction of pET28a-CHIPΔTPR, pET28a-R4-uAb and pET28a-D10-uAb has been previously described (see Chapter 2 Materials and Methods). Separately, each of the DBPs including scFv GCN4 [34], scFv 5x1 (Waraho unpublished results), scFv HAG (a gift from Andreas Pluckthun), scFv 3DX (a gift from Bruce Mayer), scFv J21 (a gift from Andreas Pluckthun) and DARPin off7 (a gift from Marc Ostermeir) were PCR amplified with a 5' *Nco*I site and 3' *Eco*RI site and ligated into pET28-DBP-uAb (**Fig. 2.2a**). Due to an internal *Eco*RI site in the monobody YS1 (a gift from Shohei Koide), overlap extension PCR was used to add the GSGSG linker and N-terminus of CHIPΔTPR to the YS1 PCR product, until reaching a unique *Bst*BI site within CHIPΔTPR. This overlap extension PCR product was then ligated between *Nco*I and *Bst*BI sites of pET28a-R4-uAb, yielding pET28a-YS1-uAb. The genes encoding scFvs D10 and J21 were also PCR amplified with a 5' *Nco*I site and 3' *Sal*I site and ligated into pET28-DBP-uAb carrying the double epitope Flag-6xHis tag, yielding the control plasmids pET28a-D10 and pET28a-J21, respectively.

For expression in eukaryotic cells, select constructs were PCR amplified from their pET28a(+) backbones using primers that introduced a Kozak sequence at the start codon as well as a 5' *Hin*dIII site and a 3' *Xba*I site. The resulting PCR products were then cloned between the *Hin*dIII and *Xba*I sites of plasmid pcDNA3, yielding pcDNA3-D10, pcDNA3-YS1-uAb and pcDNA3-off7-uAb. Due to an internal *Sal*I site, the YS1 gene was directly cloned into pcDNA3 for the construction of pcDNA3-YS1 using overlap extension to add a *Hin*dIII site and Kozak sequence to the 5' end and FLAG tag, 6x-His tag and *Xba*I site to the 3' end. The transformation

control plasmid pGFP was a gift from Pengbo Zhou.

The target substrate gpD (a gift from Andreas Pluckthun) was PCR amplified from using primers with a 5' *Nde*I site and a 3' *Hin*dIII site. DNA encoding a double tag of 6x-His-HA was created by dimerizing primers with a 5' *Nco*I overhang and a 3' *Nde*I overhang. Double ligation was performed to insert the gpD PCR product and primer dimer between *Nco*I and *Hin*dIII sites in pET28a(+), yielding plasmid pET28a-6xHis-HA-gpD. Using this backbone, the substrate JNK2 (a gift from Andreas Pluckthun) was PCR amplified using primers with a 5' *Nde*I site and a 3' *Hin*dIII site, digested and ligated to create pET28a-6xHis-HA-JNK2. For eukaryotic expression, HA-gpD was PCR amplified from the pET28a construct with a 5' and 3' *Not*I site for ligation into pCMV. The target substrate protein MBP was PCR amplified using primers that introduced a Kozak sequence at the start codon as well as 5' *Hin*dIII and a 3' *Xba*I site, and cloned in the corresponding sites of pcDNA3.

Protein expression and purification. All purified proteins were obtained from cultures of *E. coli* BL21(DE3) cells grown in 50 mL of Luria-Bertani (LB) medium. Expression was induced with 1 mM IPTG when the culture density (Abs_{600}) reached 0.6-0.8 and proceeded at 30°C for 6 h, after which cells were harvested by centrifugation at 4,000xg for 20 min at 4°C. The resulting pellets were stored at -20°C overnight. Thawed pellets were resuspended in 4 mL binding buffer (50 mM TrisHCl pH 7.9, 500 mM NaCl, 1% Tween20 and 5 mM imidazole) and lysed by sonication (30 sec three times, per 1 mL aliquot). Lysates were cleared by centrifugation at 12,000xg for 15 min at 4°C and then subjected to Ni^{2+}-affinity purification with Ni-NTA Spin Columns (Qiagen). Samples were washed with the Tris-based buffer with 60 mM imidazole before elution in 250mM imidazole buffer lacking 1% Tween20.

Enzyme-linked immunosorbent assay (ELISA). For ELISA, a previously established protocol

was used to detect binding to gpD and JNK2 [33], with slight modification. Briefly, a 96-well EIA plate was coated overnight at 4°C with 100 µL 6x-His-HA-JNK2 at 10 µg/mL in PBS or 100 µL 6x-His-HA-gpD in coating buffer (0.05 M Na_2CO_3 pH 9.6). The plate was then washed three times with 100 µL PBST (1x PBS + 0.1% Tween20) per well for 5 min with shaking and blocked with 100 µL PBS-BSA (1x PBS + 50 µg/mL BSA) per well at room temp, slowly mixing for 2 h. Following three washes as above, purified protein samples were introduced in blocking buffer as serial dilutions with 40 µL per well and incubated at room temp slowly mixing for 1 h. Three washes were used to remove non-bound protein before introducing 50 µL of anti-flag-HRP (diluted 1:10,000 in PBS) and incubating at room temp with slow mixing for 1 h. Three final washes were performed before incubation with 200 µL OPD (Sigma Fast tablets) in the dark for 30 min. The reaction was then quenched with 50 µL 3N H_2SO_4 and absorbance read at 492 nm. Notably, the ELISA procedure used here was later improved upon by using milk for blocking instead of BSA, which dramatically reduced the binding of negative controls such as CHIPΔTPR (see Chapter 2 Materials & Methods).

***In vitro* ubiquitination assay.** Ubiquitination assays were performed as previously described [48] in the presence of 0.1 µM purified human recombinant UBE1 (Boston Biochem), 2 µM human recombinant UbcH5α/UBE2D1 (Boston Biochem), 3 µM uAb (or equivalent control protein), 3 µM 6xHIS-HA-gpD or 3 µM inactivated JNK2 (Invitrogen), 50 µM human recombinant ubiquitin (Boston Biochem), 4 mM ATP and 1 mM DTT in 20 mM MOPS, 100 mM KCl, 5 mM $MgCl_2$, pH 7.2. Alternatively for Coomassie evaluation, reactions were performed as for MS analysis, with 0.1 µM UBE1, 20 µM UbcH5α, 20 µM D10-uAb, 20 µM gpD and 500 µM ubiquitin. Reactions were carried out at 37°C for 2 h (unless otherwise noted) and stopped by boiling in 2x Laemmli loading buffer for analysis by immunoblotting.

Cell culture and transfection. HEK293T cells were cultured in DMEM with 10% heat inactivated FBS and 1% antibiotic-antimycotic (Cellgro) at 37°C with 5% CO_2. Cells were transfected at 60-80% confluency with jetPRIME® (Polyplus Transfection) using 2 μg total DNA per well in a 6-well plate with a 1:2 jetPRIME® ratio and at 4 h post-transfection the growth media was refreshed. At 24 h post-transfection, cells were harvested in PBS and frozen at -20°C until analyzed by immunoblotting. Thawed cells were lysed in NP40 lysis buffer (150 mM NaCl, 1% Nonidet P-40 and 50 mM TrisHCl, pH 7.4) by pipetting and mixing at 4°C for 30 min, followed by removal of the insoluble fraction at 18,000xg at 4°C for 20 min. Insoluble pellets were then washed with 50mM TrisHCl and 1mM EDTA, pH 8 followed by solubilization in an equal volume of 2% SDS in PBS by boiling for 10 min. Cooled samples were centrifuged at room temp for 10 min at 13,200 RPM to remove remaining cellular debris. Both soluble and insoluble fractions were boiled in 2x Laemmli sample buffer for analysis by immunoblotting. Soluble fraction lysates were normalized using a detergent compatible total protein assay (Bio-Rad) and 10 μg total protein was loaded with equivolume insoluble fractions for immunoblotting comparison.

Pull-down assays. Thawed cells were lysed as above; then clarified lysates were normalized by a detergent compatible total protein assay (Bio-Rad) and diluted to contain 0.75 mg/mL of protein with 10 mM imidazole to reduce non-specific binding. 200 μL of diluted lysates were incubated with 30 μL of Ni-NTA magnetic agarose beads (Qiagen, 5% solution) mixing at 4°C for 1-2 h and washed with 20 mM imidazole in lysis buffer. Bound proteins were eluted with 250 mM imidazole (50 mM TrisHCl pH 7.9 and 50 mM NaCl) and boiled in 2x Laemmli sample buffer for analysis by immunoblotting.

Protein analysis. SDS-PAGE and immunoblotting of proteins was performed according to

standard procedures. BioRad Coomassie G-250 stain was used to visualize proteins in SDS-PAGE (BioRad, Mini-PROTEAN® TGX). The following primary antibodies were utilized for immunoblotting: rabbit anti-HA (Sigma, H6908), mouse anti-JNK2 (MBLI, 0301), mouse anti-ubiquitin (Millipore, P4D1-A11), rabbit anti-Lys48 (Millipore, Apu2), rabbit anti-6x-His-HRP (Abcam, ab1187), mouse anti-GAPDH (Millipore, 6C5), mouse anti-FLAG®-HRP (Abcam, ab49763), mouse anti-Hsp70 (Enzo Life Sciences, C92F3A), mouse anti-GFP (Roche, clones 7.1 and 13.1), mouse anti-MBP-HRP (NEB, E8038). Secondary antibodies goat anti-rabbit IgG (H+L) and anti-mouse IgG (H+L) with HRP conjugation (Promega) were utilized as needed. For ELISA detection, mouse anti-FLAG®-HRP (Abcam, ab49763) was used.

Acknowledgements

We thank Cam Patterson, Pierre Martineau, Bruce Mayer, Andreas Plückthun, Pengbo Zhou, Marc Ostermeir, and Shohei Koide for kindly providing plasmids encoding genes used in this study. This material is based upon work supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Institutes of Health Chemical Biology Training Grant T32 GM008500 (fellowship to A.D.P.) and the National Science Foundation GK-12 DGE-0841291 (fellowship to A.D.P.).

CHAPTER 4

RATIONAL DESIGN OF UBIQUIBODIES FOR ENHANCED ACTIVITY

Introduction

The ubiquitin ligase CHIP (carboxyl terminus of Hsc70-interacting protein) was first discovered due to its N-terminal tetratricopeptide repeat (TPR) domain that allows it to bind and regulate heat shock proteins (Hsp) which act as chaperones for unfolded or misfolded proteins [46]. It was later realized that this 35 kDa cytoplasmic protein also contains a C-terminal U-box domain, which binds ubiquitin conjugating enzymes (E2s) and mediates ubiquitin ligase activity [74]. The U-box is a 75 amino acid domain first identified in yeast and is considered a non-canonical RING domain, thus classifying proteins containing it as E3 ubiquitin ligases [75]. In combining the functions of the N- and C-terminal domains, CHIP has the ability to assist in both chaperone-associated folding and the degradation of chaperone substrates, giving it a unique role in protein quality control [74]. Indeed, it has been shown that CHIP contains intrinsic chaperone functions in collaboration with Hsp70 and independently [47, 76].

To date, the ubiquitin ligase activity of CHIP has been shown to target a myriad of proteins for proteasomal degradation. These include chaperone-bound substrates (e.g. cystic-fibrosis transmembrane-conductance regulator [77], glucocorticoid receptor [78], and β-amyloid [79]), the chaperones themselves (e.g. Hsc70, Hsp70 and Hsp90 [67]), and directly bound target proteins (e.g. apoptosis signal-regulating kinase 1 [80], Runx1 [81], and Smad1 [82]). This substrate diversity also reveals CHIP's functionality in multiple sub-cellular compartments including the endoplasmic reticulum, nucleus and cytoplasm [75]. It has been shown that CHIP sequentially ubiquitinates chaperone-bound substrates before the chaperone protein itself [67]. As well, CHIP interacts with a variety of E2s, including the UbcH5 family in the formation of

diverse polyubiquitin chains, including Lys48-linked chains, and with the E2 heterodimer Ubc13-Uev1a in the formation of Lys63-linked chains [59, 75].

Furthermore, the crystal structure of CHIP has been solved in complex with the C-terminal peptide of Hsp90α , the E2 heterodimer Ubc13-Uev1a and the E2 UbcH5α [45, 50, 59]. Zhang et al. found that full-length murine CHIP co-crystallized with the human Hsp90α C-terminal peptide forms an asymmetric homodimer (**Fig. 4.1**). The main difference between the two protomers of CHIP in the asymmetric homodimer is the intrinsic flexible helical linker (α7) which takes on a "straight" (**Fig. 4.1a**) or "bent" (**Fig. 4.1b**) conformation. The dimer has interfaces of hydrophobic packing between the U-boxes and α-helix 7 (**Fig. 4.1c**), supporting the finding that the flexible helical linker is required for dimerization and functionality [83].

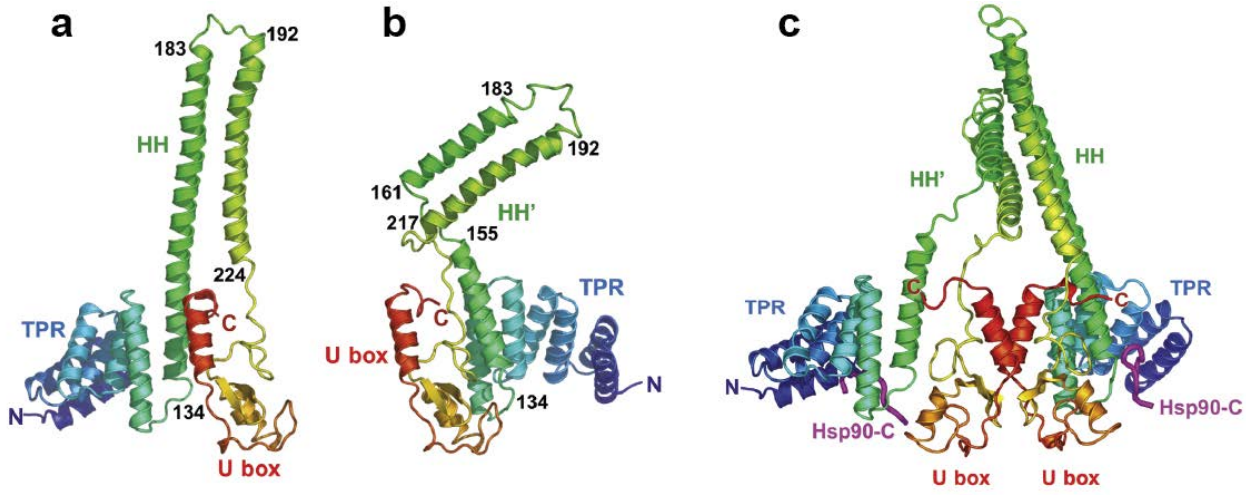


Figure 4.1 Crystal structure of CHIP (taken from Zhang et al, [45]). Secondary structure cartoon of the elongated CHIP protomer (a) compared to the bent conformation (b). The protein is rainbow colored (blue to red) from the N- to C-terminus, showing the tetratricopeptide repeat domain (TPR), helical hairpin linker (HH), and the U-box domain. (c) The asymmetric CHIP homodimer, with the co-crystallized Hsp90α C-terminal decapeptide bound to each TPR domain shown in magenta. Images were produced by [45] using MacPyMOL.

In studying the asymmetric dimer, they realized that CHIP was only able to bind its E2, in the straight conformation because in the bent conformation the U-box is blocked by the TPR domain (**Fig. 4.1c**). Functionally, this means that only half of the CHIP dimer is available for E2

activity, which may help facilitate the formation of uniform polyubiquitin chains (i.e. not mixed linkages) [45]. They also found that the U-box domain interacts solely with Ubc13 which then recruits and binds Uev1a to form the heterodimeric E2. Upon co-crystallization of CHIP's U-box domain with the E2 UbcH5α, Xu et al. found that both Ubc13 and UbcH5 interface with the same surface of the U-box domain of CHIP (**Fig. 4.2a** (Ubc13) and **4.2b** (UbcH5α)).

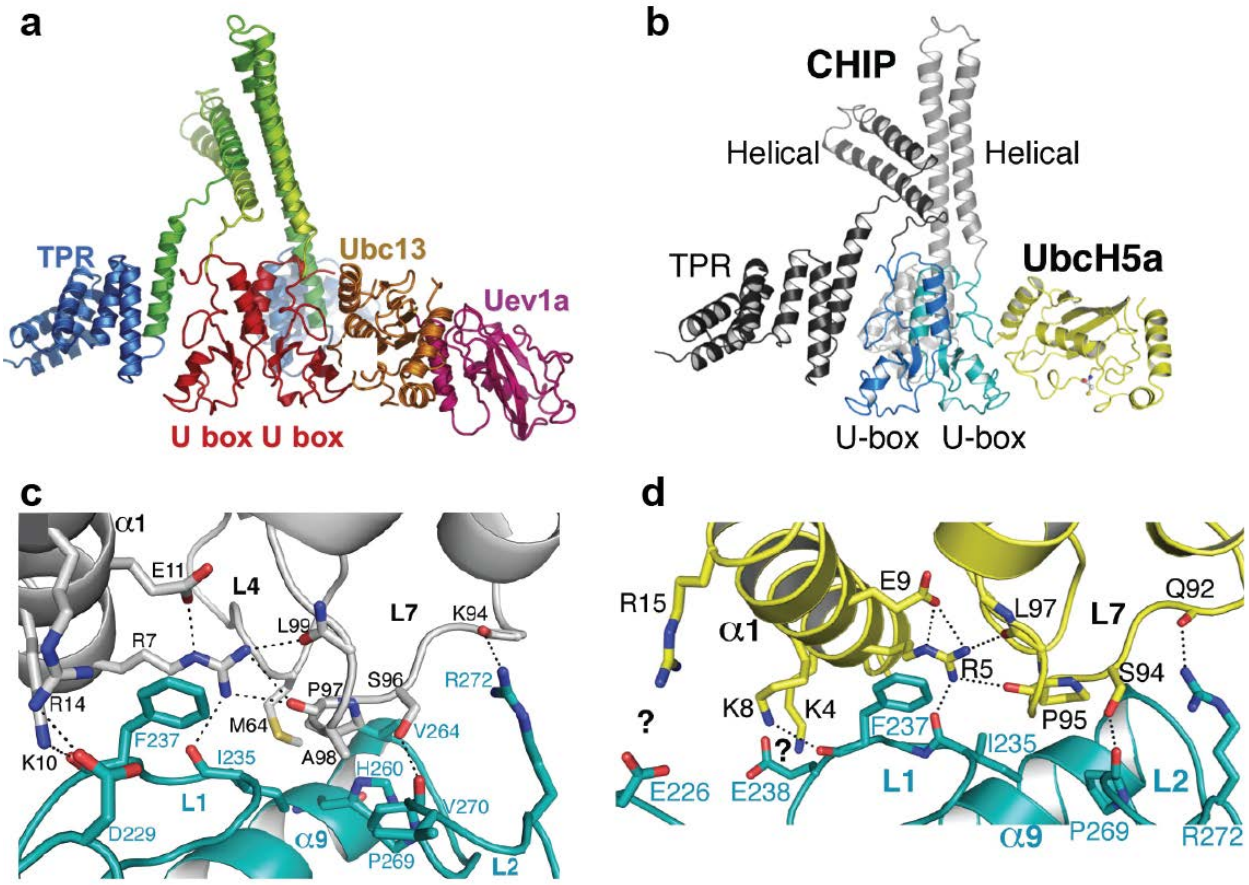


Figure 4.2 CHIP co-crystal structures with E2s. Models of CHIP co-crystallized with E2s, Ubc13-Uev1a (a, [45]) or UbcH5α (b, [59]). E2s Ubc13-Uev1a and UbcH5α were co-crystallized with the U-box domain of CHIP, then U-box dimers from the co-crystal structures were superimposed on the full-length CHIP dimer crystal structure to create the models (a) and (b) respectively. Detailed interactions between the S-P-A motif of Ubc13 (c) and UbcH5α (d) and CHIP's U-box. Likely hydrogen bonds or salt bridges are shown as dotted lines. Images taken from [59] were made in Pymol.

Notably, both UbcH5 and Ubc13 E2s have a conserved Ser-Pro-Ala motif in the binding pocket which makes a hydrogen bond with the carbonyl group of P269 and has van der Waals interactions with H260, V264 and V270 in the U-box domain of CHIP (**Fig. 4.2c** and **4.2d**) [59]. However, differences in the E2 binding pockets informed mutagenesis studies to determine which residues are essential for ubiquitination activity with each E2 [59].

Clearly, the structural and functional information about CHIP made it an ideal ubiquitin ligase to use in ubiquibody development. Yet there were still features which could potentially be improved upon by rational design. Specifically, we hypothesized that we could refine the E2 specificity of ubiquibodies using the aforementioned mutagenesis studies. This is important because E2s play a significant role in directing the polyUb chain linkage formation which target substrates to the proteasome [14]. Secondly, we could use rational design to modulate the conformational flexibility of ubiquibodies required for substrate ubiquitination [48]. Finally, we sought to reduce the autoubiquitination of ubiquibodies which could increase protein half-life.

Results

Refining E2 ubiquitin conjugating enzyme preference. CHIP's diversity of E2 interactions may consequently diversify the polyUb chain formation on target proteins, as E2s often determine the polyUb chain linkage [14]. Thus in an attempt to reduce non-proteasomal polyUb linkages, we utilized studies by Xu et al. of CHIP:E2 interactions which revealed that the K234A point mutant reduced Ubc13-Uev1a interaction, which is known to create K63-linked chains, while maintaining UbcH5 binding and activity (**Fig. 4.3a**) [59]. This point mutation was made in R4-uAb and tested for *in vitro* functionality with UbcH5α and the target β-gal (**Fig. 4.3b**). While ubiquitination of β-gal was still detectable, it was notably reduced compared to wild-type R4-uAb.

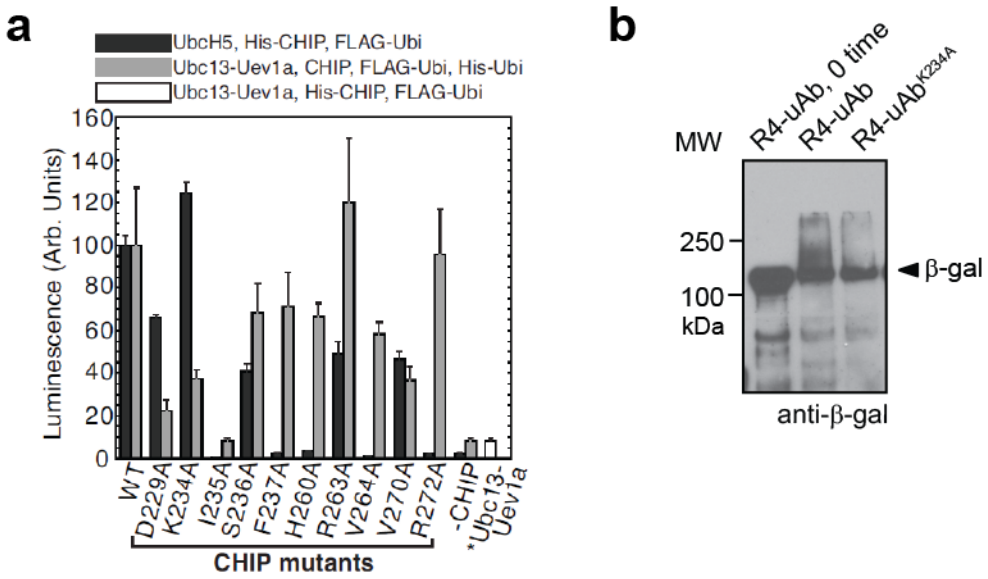


Figure 4.3 CHIP point mutants that specify E2 interactions. (a) Taken from Xu et al., i*n vitro* ubiquitination assays carried out in ELISA format detecting His-CHIP autoubiquitination with UbcH5β or free polyUb chains formed with Ubc13-Uev1a and untagged CHIP [59]. (b) *In vitro* ubiquitination assays carried out for 2 h with the E2, UbcH5α, substrate β-gal and E3s, R4-uAb or R4-uAb^{K234A} as indicated. Ubiquitination assay mixtures boiled prior to incubation were used as controls to identify unmodified proteins (0 time). An equal amount of total protein was loaded per lane and immunoblots were probed with antibodies for β-gal.

We also utilized the U-box point mutations, I235A and R272A, which were identified to interfere with UbcH5 binding (**Fig. 4.3a**). These served as negative controls to show the dependence of target ubiquitination on R4-uAb:E2 interactions (**Fig. 4.4**). Specifically, R4-uAb was compared to R4-uAb^{I235A} using *in vitro* ubiquitination reactions evaluated over time and showed no polyUb even after 2 h at 37°C, as evaluated with antibodies detecting both β-gal and ubiquitin (**Fig. 4.4a**). Similarly, R4-uAb^{R272A} was tested for ubiquitination of β-gal compared to wild-type R4-uAb and the controls scFv13-R4, CHIPΔTPR, and the non-specific D10-uAb (**Fig. 4.4b**). This confirmed the requirement for both E2 and substrate interaction with the uAb for targeted ubiquitination to proceed. While previous works utilized H260Q or P269A as CHIP U-box point mutants to disrupt E2 binding, we found that both point mutations significantly reduced R4-uAb expression in *E. coli* and therefore results were inconclusive in purified *in vitro*

ubiquitination assays (data not shown) [4, 67].

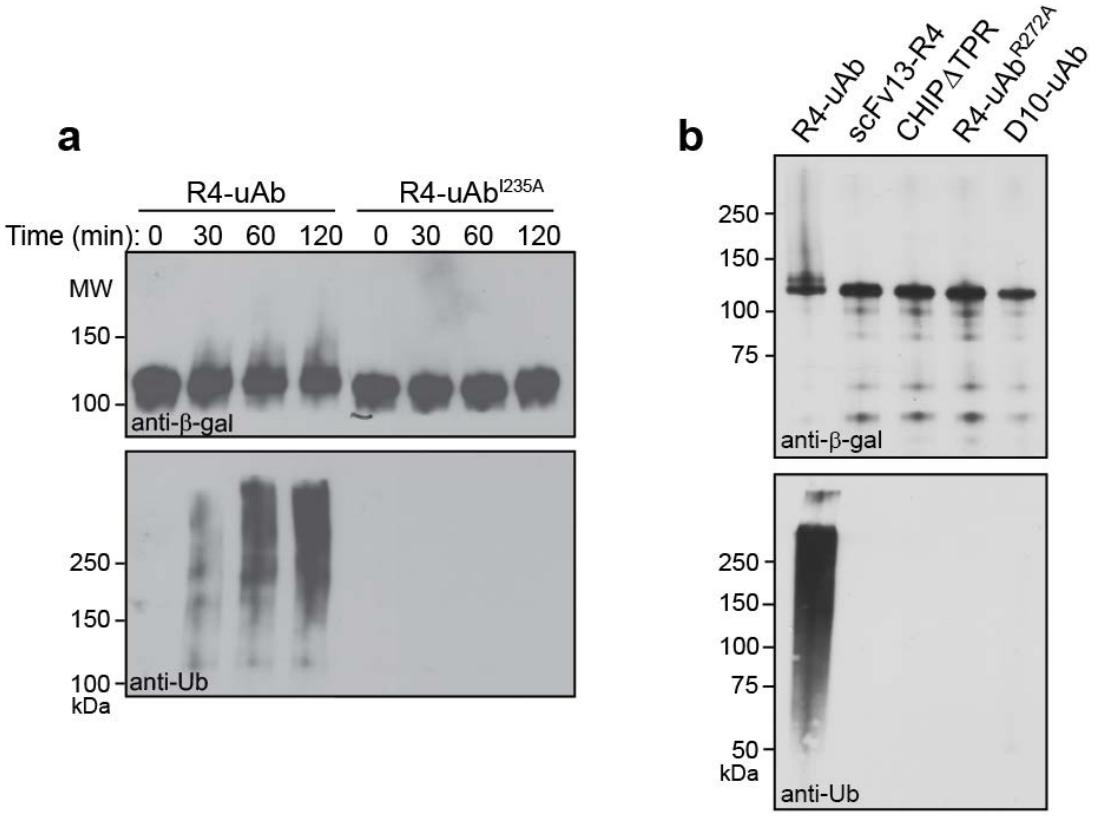


Figure 4.4 U-box point mutants for inhibiting E2 interactions. *In vitro* ubiquitination of β-gal was carried out with R4-uAb, R4-uAb^{I235A} (a), scFv13-R4, CHIPΔTPR, D10-uAb, or R4-uAb^{R272A} (b) as E3s. An equivalent amount of total protein was added to each lane. Immunoblots were probed with anti-β-gal and anti-ubiquitin antibodies. Molecular weights of the marker bands (MW) are indicated.

All R4-uAb point mutants known to interfere with E2 interactions were next tested in HEK293T cells to determine if E2 inhibition would reduce or eliminate β-gal knockdown (**Fig. 4.5**). On two separate occasions, none of the R4-uAb U-box mutants reduced β-gal knockdown as would be predicted. In fact, many of the point mutations appeared to knockdown β-gal levels as well as, or better than wild-type R4-uAb (**Fig. 4.5** and data not shown). It was also evident that all U-box mutant ubiquibodies had reduced expression in HEK293T (**Fig. 4.5**). However, both replicates of testing U-box point mutants were performed with the high level of pcDNA3-β-gal

transfection where knockdown was convoluted by sequestration to the insoluble fraction (see **Fig. 2.11** and **Fig. 2.12**). As noted in chapter 2, the reduced levels of β-gal expression seen could be largely due to sub-cellular partitioning, rather than proteasomal degradation. Interestingly, $CHIP^{H260Q}$ has previously been shown to preferentially sequester a natural substrate, iNOS (inducible nitric oxide synthase), to a detergent-resistant fraction [84]. This suggests that the U-box activity mutants may show an enhanced loss of soluble β-gal compared to wild-type R4-uAb by sequestering more β-gal into the detergent-resistant fraction rather than having more efficient UPP degradation. In fact, preliminary proof was shown in Figure 2.13 where HEK293T cells were co-transfected with ten-fold less pcDNA3-β-gal and $R4\text{-}uAb^{R272A}$, yet the insoluble fraction of cells still revealed a small amount of β-gal. Thus, further characterization of point mutants which interfere with E2 binding will need to be made in eukaryotic cells to determine the nature of β-gal knockdown detected. Minimally, ubiquibody point mutants should be re-tested at the lower β-gal transfection level used to validate proteasomal knockdown in chapter 2. Further studies could be an alternative future direction for ubiquibody research and will be elaborated upon in chapter 5.

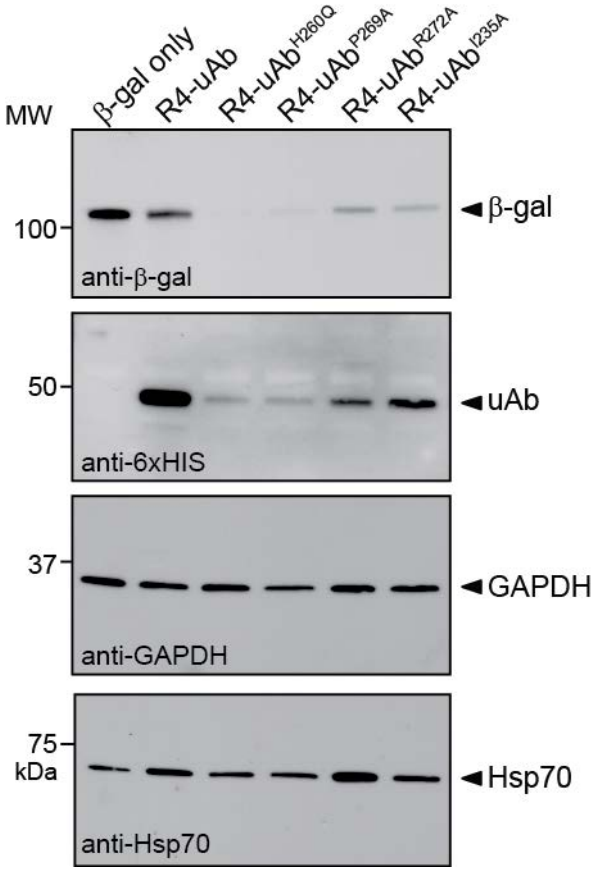


Figure 4.5 Ubiquibody U-box mutants in HEK293T cells. Immunoblots of soluble extracts prepared from HEK293T cells transfected with 0.5 μg pcDNA3-β-gal alone (β-gal only) or co-transfected with pcDNA3-R4-uAb, pcDNA3-R4-uAb^{mutant}, or pcDNA3-D10-uAb. An equivalent amount of total protein was loaded in each lane. Blots were probed with antibodies specific for β-gal, 6xHis, GAPDH and Hsp70 as indicated.

Enhancing ubiquitin ligase flexibility. Ubiquitin ligase flexibility is necessary for substrate ubiquitination due to the 'macromolecular juggling' required amidst E2 and substrate interactions [85]. In CHIP, the α-helical domain, between the two binding domains, is required to maintain conformational flexibility for active ubiquitination [48]. Our early tests of solubility in *E. coli* revealed that the addition of a short Gly-Ser-Gly-Ser-Gly (GSGSG) linker significantly improved ubiquibody expression (**Fig. 4.6a**). Thus as we began testing intrabodies with reduced stability (e.g. scFvs 5x1, C4 and NAC32), we introduced longer linkers to determine if these could further improve uAb expression (**Fig. 4.6b**). Specifically we tested the flexible Gly-Ser

linkers commonly found in scFvs, $(Gly_4Ser)_3$ and $(Gly_4Ser)_5$. The only discernible improvement of solubility was with NAC32-uAb which was found to be significantly degraded, but in which longer linkers did increase overall expression (**Fig. 4.6b**). Additionally, we tested purified D10-uAbs with the longer linkers for *in vitro* ubiquitination, where previously we'd only been able to detect monoubiquitination of the small substrate, gpD by immunoblotting. We hypothesized that polyUb of gpD could be enhanced with greater uAb flexibility (**Fig. 4.6c**). However, no improvement of gpD ubiquitination was detectable by immunoblotting.

While general improvements of stability or ubiquitination activity were not evident with *in vitro* ubiquitination reactions, longer linkers may prove especially valuable with the continued use of smaller format DBPs. From the asymmetric CHIP dimer crystal structure, the TPR domain effectively blocks the U-box domain from E2 binding in the bent conformation, which may prove significant in the ubiquitination of substrates [48]. Additionally, it's been shown that the natural helical linker in CHIP is necessary for dimerization and ubiquitination activity [83], so an additional flexible linker may be helpful in accommodating non-natural DBPs while maintaining the dimer interface. Alternatively, too much flexibility may hinder essential substrate:uAb:E2 interactions, (e.g. U-box domain blocking) required for E2 cycling and polyUb chain formation.

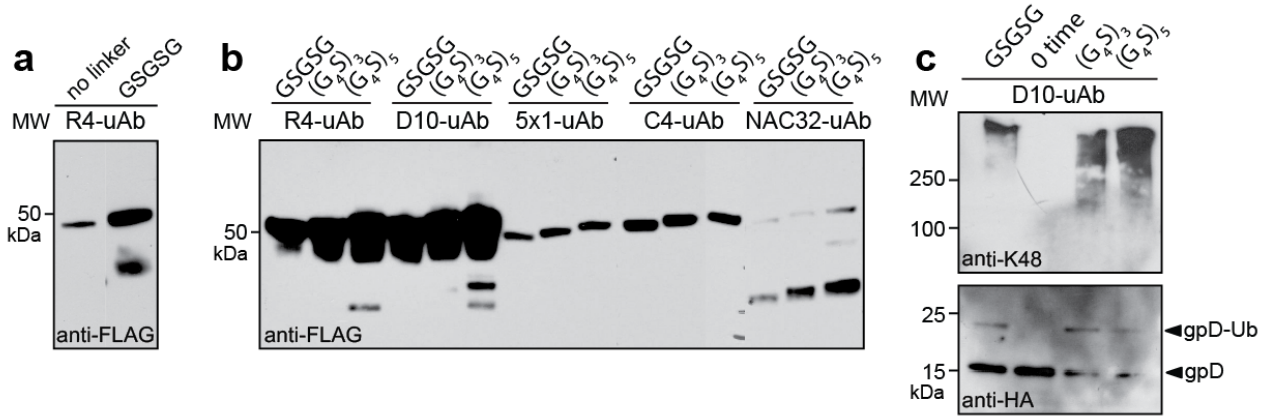


Figure 4.6 Ubiquibody linkers for enhanced flexibility. Western blot analysis of cell lysates derived from *E. coli* strain BL21(DE3) comparing the expression of ubiquibodies with (a) scFv13-R4 fused to CHIPΔTPR with and without the GSGSG linker, and (b) scFv13-R4, D10, scFv5x1, C4, and NAC32 with GSGSG, $(G_4S)_3$ or $(G_4S)_5$ linkers. Lysates were normalized by total protein and immunoblots probed with anti-FLAG® antibodies. (c) *In vitro* ubiquitination of gpD was carried out for 2 h with D10-uAbs with the linkers GSGSG, $(G_4S)_3$ or $(G_4S)_5$. A ubiquitination assay mixture with the D10-uAb (GSGSG) was boiled prior to incubation as a control to identify unmodified proteins (0 time). An equivalent amount of total protein was added to each lane and immunoblots were probed with anti-HA and anti-K48 linked polyUb antibodies as indicated.

Reducing autoubiquitination of ubiquibodies. Another design aspect worth considering is the known autoubiquitination of CHIP [48], which could diminish protein half-life. However, even though polyubiquitination of ubiquibodies was consistently observed *in vitro* (**Fig. 4.7a**), stable accumulation of ubiquibodies was observed in mammalian cells, with expression levels increasing in a plasmid dose-dependent fashion (see **Fig. 2.13** and **3.7**). Nonetheless, it might be possible to eliminate autoubiquitination by mutating acceptor lysines in CHIPΔTPR without adversely affecting the ubiquibodies' solubility, protein-protein interactions or catalytic activity. Notably, it has been shown that the RING E3, Mdm2 promotes autoubiquitination via ubiquitination of a specific lysine residue, in contrast to most site specificity studies which have revealed a lack of lysine specificity [26]. Thus, we sought to determine if any single lysine residue within CHIPΔTPR could be attributed site specific autoubiquitination. We choose only to

look at lysine residues within CHIPΔTPR with the desire to avoid adding an additional design constraint to DBPs and to avoid interfering with substrate recognition. We began by performing an alanine screen in R4-uAb of each lysine residue in the CHIPΔTPR domain, of which there are eleven. Each point mutant was then expressed in the *E. coli* strain BL21(DE3) (**Fig. 4.7b**), purified and tested for the ability to ubiquitinate the target β-gal and autoubiquitinate (**Fig. 4.7c**). Remarkably, one clone, R4-uAb^{K287A} showed decreased autoubiquitination.

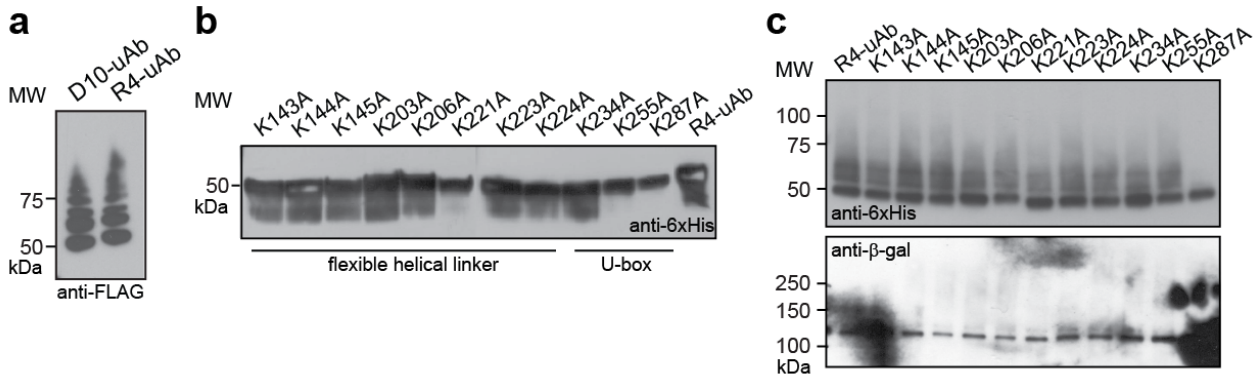


Figure 4.7 Alanine screening of lysine residues in CHIPΔTPR. (a) *In vitro* ubiquitination assays with gpD as the target protein and D10-uAb or R4-uAb as the E3s. Autoubiquitination was evident in laddering to higher molecular weight species as detected with anti-FLAG® antibodies. (b) Western blot analysis of cell lysates derived from *E. coli* strain BL21(DE3) expressing R4-uAb with Lys to Ala mutations made as noted. An equivalent amount of total protein was loaded in each lane and anti-6xHis antibodies were used to detect the expressed proteins. The structural domains in CHIPΔTPR of the mutations are annotated below. (c) *In vitro* ubiquitination assays with Lys to Ala R4-uAb mutants from (b) after purification using Ni^{2+} chromatography. Assays included β-gal as the substrate protein and used UbcH5α as the E2. An equivalent amount of protein was loaded per lane and immunoblots were probed with anti-6xHis and anti-β-gal antibodies as indicated.

Next the R4-uAb^{K287A} mutant was tested in HEK293T cells to determine if reduced autoubiquitination would improve upon wild-type R4-uAb knockdown of β-gal (**Fig. 4.8a** and **b**). Preliminary studies showed that R4-uAb^{K287A} was expressed at a lower level in HEK293T cells than wild-type and thus the more conservative K287R mutation was also made and found to recover some of the uAb expression (**Fig. 4.8a**). Unfortunately both reduced solubility (due to

misfolding) and autoubiquitination would lead to proteasomal degradation of the ubiquibody, making decoupling these two outcomes difficult. Furthermore, both R4-uAb^{K287A} and R4-uAb^{K287R} gave mixed results with regards to β-gal silencing in HEK293T cells (**Fig. 4.8a** and **b**) Thus, preliminary studies towards reducing autoubiquitination of uAbs by single lysine mutations proved inconclusive in the presence of the natural ubiquitin proteasome pathway.

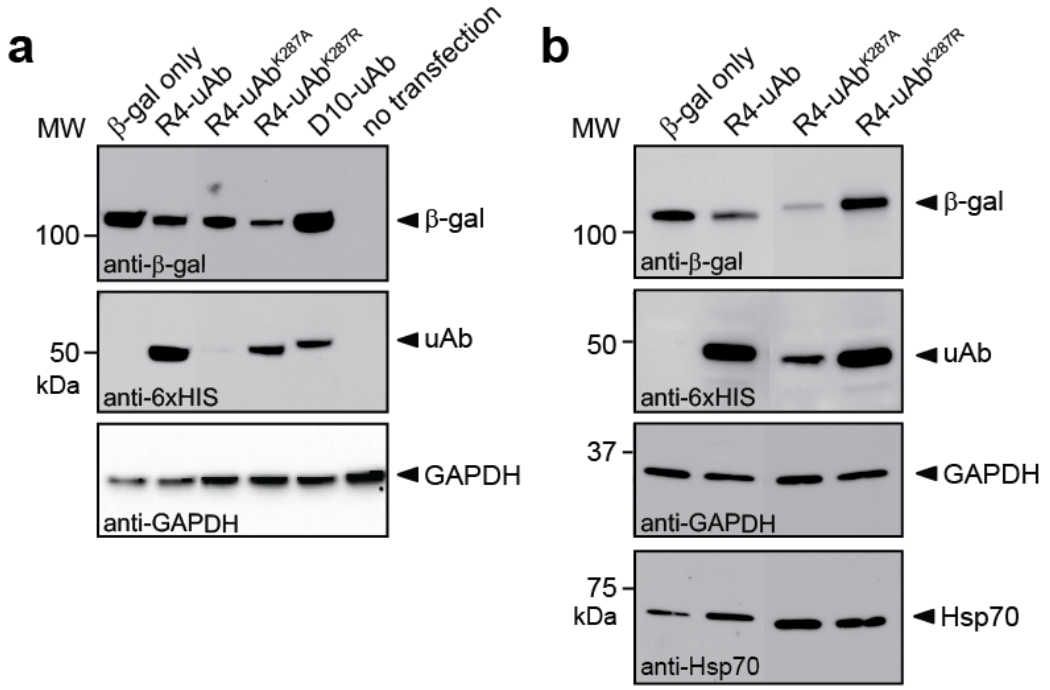


Figure 4.8 Evaluating uAb autoubiquitination in HEK293T cells. (a) Immunoblots of soluble extracts prepared from HEK293T cells transfected with 0.5 μg pcDNA3-β-gal alone (β-gal only) or co-transfected with pcDNA3-R4-uAb, pcDNA3-R4-uAb^{K287A}, pcDNA3-R4-uAb^{K287R}, or pcDNA3-D10-uAb. An equivalent amount of total protein was loaded in each lane. Blots were probed with antibodies specific for β-gal, 6x-His, GAPDH and Hsp70 as indicated. (b) Replicate experiment of (a).

Discussion

At this point, preliminary tests have been performed with engineered ubiquibodies towards specifying E2 interactions, improving flexibility and reducing autoubiquitination. Further testing of the U-box K234A mutant, which reduces interactions with Ubc13, could prove

valuable in deciphering specifically how β-gal knockdown is occurring (i.e. K63-linked versus K48-linked polyUb chains). Similarly U-box activity point mutants H260Q, P269A, I235A and R272A in R4-uAb should be tested at lower β-gal transfection levels with insoluble fraction evaluation, to determine if detergent-resistant partitioning is common amongst all mutants. Notably, when full-length CHIP targets endogenous substrates to the detergent-resistant fraction, it requires interactions with heat shock proteins [84]. However, the partitioning seen with ubiquibodies may be due to the formation of K63-linked chains which can be recognized by HDAC6 and sent to the juxtanuclear aggresome [19]. Thus, it would be informative to perform a pull-down with 6xHis-tagged ubiquibodies in mammalian cells to determine what endogenous proteins may be interacting with the ubiquibodies (i.e. utilizing antibodies against Hsp70 and endogenous CHIP).

With regards to autoubiquitination, a more careful analysis of protein half-life could be performed using the ribosomal inhibitor cycloheximide. This could elucidate if point mutations actually reduce autoubiquitination and degradation in mammalian cell lines. However, as previously mentioned, decoupling degradation due to misfolding versus autoubiquitination may prove more difficult. Finally, continued rational design of ubiquibodies would be greatly aided by the crystal structure of uAbs, and the co-crystal structures of uAb:E2 and uAb:substrate interactions. If ubiquibodies dimerize with straight and bent conformations, it may be necessary for DBPs to similarly block the U-box domain in the bent conformation. As such, this could put a size limit on DBPs and would require the substrate to be able to bind the uAb in the bent conformation. Thus, a crystal structure could inform the utility of flexible linkers and would confirm whether ubiquibodies function similar to wild-type CHIP in an asymmetric homodimer conformation.

Materials and Methods

Plasmid construction. Site directed mutagenesis to create all point mutations in pET28a(+)-R4-uAb was performed using QuikChange II (Stratagene) techniques with modifications. Briefly, primers were designed to be ~30 base pairs with the desired mutation in the middle of the primer, a $T_m \geq 78$°C, a GC content greater than 40%, and terminating in one or more G or C. T_m was calculated using the following formula: $T_m = 81.5 + 0.41(\%\ GC) - (675/N) - \%$ mismatch, where N is the primer length in bases and % GC (% guanidine and cytosine content) and % mismatch are whole numbers. PCR reactions were set-up with 50 ng of each primer, 5-20 ng of template DNA, 0.2 mM each dNTP, 0.5 μL Phusion HF (2 U/μL), 4 μL 5xGC buffer and water to a final 20 μL volume. PCRs were performed using touchdown annealing temperatures from 70°C to 55°C and extension at 7 °C giving Phusion HF 30 sec per kilobase of plasmid. PCRs were then digested with 1 μL of *Dpn*I restriction enzyme (20 U/μL) at 37°C for 1 h. Reactions were then desalted for 20 min and 5 μL transformed into electrocompetant DH5α *E. coli*. After recovery, cells were grown overnight on solid growth media containing kanamycin, single colonies were isolated and plasmid DNA was sent to sequencing to confirm single point mutations.

To create longer linkers in pET28a-uAb, the forward primers listed in Table 4.1 were used in nested PCRs to extend the 5' end of CHIPΔTPR to add the appropriate sequences, $(Gly_4Ser)_3$ or $(Gly_4Ser)_5$ respectively with a 5' terminal *Eco*RI site for cloning back into pET28a-uAb. The reverse primer was used to amplify the 3' end of CHIPΔTPR and includes the 3' *Sal*I site prior to the double epitope tags in pET28a-uAb (see **Fig. 2.2a**).

Table 4.1 Primers used to create $(Gly_4Ser)_n$ linkers

Primer Name	Sequence	T_m (°C)
$(Gly_4Ser)_n$ innermost primer fwd	ggt gga ggc agc ggt ggc gga ggc tca gga ggc ggt ggc tcccggctgaacttcggggac	58.7
$(Gly_4Ser)_5$ outer primer fwd	ct cag GAATTC ggt gga ggt ggc agt ggc gga ggt ggc tct ggc ggt gga ggc agc ggt g	59.6
$(Gly_4Ser)_3$ outer primer fwd	cat cag GAATTC ggc ggt gga ggc agc ggt g	59.6
SalI-hCHIPdTPR-rev	ctg atg GTC GAC gta atc ctc cac cca gcc att c	59

Additional intracellular scFvs including scFv 6E [86], scFv C4 [87], and scFv NAC32 [88] were synthesized by GenScript based on NCBI sequences with a 5' *Nco*I site and 3' *Eco*RI site in pUC57 plasmids. These were then digested and ligated into pET28-DBP-uAb (**Fig. 2.2a**). For expression in eukaryotic cells, select constructs were PCR amplified from their pET28a(+) backbones using primers that introduced a Kozak sequence at the start codon as well as a 5' *Hin*dIII site and a 3' *Xba*I site. The resulting PCR products were then cloned between the *Hin*dIII and *Xba*I sites of plasmid pcDNA3.

Protein expression and purification. Purified proteins were obtained from cultures of *E. coli* BL21(DE3) cells grown in 50 mL of Luria-Bertani (LB) medium. Expression was induced with 0.1 mM IPTG when the culture density (Abs_{600}) reached 0.6-0.8 and proceeded at 30°C for 6 h, after which cells were harvested by centrifugation at 4,000xg for 20 min at 4°C. The resulting pellets were stored at -20°C overnight. Thawed pellets were resuspended in 3 mL binding buffer (50 mM TrisHCl pH 7.9, 500 mM NaCl, 1% Tween20 and 20 mM imidazole) and lysed by sonication (30 sec three times, per 1 mL aliquot). Lysates were cleared by centrifugation at 12,000xg for 15 min at 4°C and then subjected to Ni^{2+}-affinity purification with Ni-NTA Spin Columns (Qiagen). Samples were washed with the Tris-based buffer with 60 mM imidazole before elution in 250 mM imidazole buffer lacking 1% Tween20.

***In vitro* ubiquitination assay.** Ubiquitination assays were performed as previously described

[48] in the presence of 0.1 μM purified human recombinant UBE1 (Boston Biochem), 2 μM human recombinant UbcH5α/UBE2D1 (Boston Biochem), 3 μM uAb (or equivalent control protein), 3 μM *E. coli* β-gal (Sigma) or 3 μM 6xHIS-HA-gpD, 50 μM human recombinant ubiquitin (Boston Biochem), 4 mM ATP and 1 mM DTT in 20 mM MOPS, 100 mM KCl, 5 mM $MgCl_2$, pH 7.2. Reactions were carried out at 37°C for 2 h (unless otherwise noted) and stopped by boiling in 2x Laemmli loading buffer for analysis by immunoblotting.

Cell culture and transfection. HEK293T cells were cultured in DMEM with 10% heat inactivated FBS and 1% antibiotic-antimycotic (Cellgro) at 37°C with 5% CO_2. Cells were transfected at 60-80% confluency with jetPRIME® (Polyplus Transfection) using 2 μg total DNA per well in a 6-well plate with a 1:2 jetPRIME® ratio and at 4 h post-transfection the growth media was refreshed. At 24 h post-transfection, cells were harvested in PBS and frozen at -20°C until analyzed by immunoblotting. Thawed cells were lysed in NP40 lysis buffer (150 mM NaCl, 1% Nonidet P-40 and 50 mM TrisHCl, pH 7.4) by pipetting and mixing at 4°C for 30 min, followed by removal of the insoluble fraction at 18,000xg at 4°C for 20 min. Soluble fraction lysates were normalized using a detergent compatible total protein assay (Bio-Rad), boiled in 2x Laemmli sample buffer, and 10 μg total protein was loaded on SDS-PAGE for immunoblotting.

Protein analysis. SDS-PAGE and immunoblotting of proteins was performed according to standard procedures. BioRad Coomassie G-250 stain was used to visualize proteins in SDS-PAGE (BioRad, Mini-PROTEAN® TGX). The following primary antibodies were utilized for immunoblotting: rabbit anti-β-gal (Abcam, ab616), mouse anti-ubiquitin (Millipore, P4D1-A11), rabbit anti-Lys48 (Millipore, Apu2), rabbit anti-6x-His-HRP (Abcam, ab1187), mouse anti-GAPDH (Millipore, 6C5), mouse anti-FLAG®-HRP (Sigma, M2), mouse anti-Hsp70 (Enzo Life Sciences, C92F3A), rabbit anti-HA (Sigma, H6908). Secondary antibodies goat anti-rabbit IgG

(H+L) and anti-mouse IgG (H+L) with HRP conjugation (Promega) were utilized as needed.

Acknowledgements

I'd like to thank Emily Buirkle, CBE '12, for assistance in testing the ubiquibody linkers as well as improving upon purifications of low expressing uAbs. I'd also like to thank CBE graduate student Sarai Meyer for assistance in cloning the CHIPΔTPR alanine mutants during her rotation in our lab. We thank Cam Patterson, Pierre Martineau, and Andreas Plückthun for kindly providing plasmids encoding genes used in this study. This material is based upon work supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Institutes of Health Chemical Biology Training Grant T32 GM008500 (fellowship to A.D.P.) and the National Science Foundation GK-12 DGE-0841291 (fellowship to A.D.P.).

CHAPTER 5

FUTURE DIRECTIONS OF UBIQUIBODY TECHNOLOGY

Introduction

Ubiquitination signaling has diverse roles in maintaining cellular homeostasis. Its activity varies from cellular defense against bacteria [89], to regulating endocytosis, mitosis and inflammatory responses, to aiding in DNA damage repair and targeting misfolded proteins for degradation [16]. Not unsurprisingly, this complex ubiquitin code is still not fully understood. However, as we seek to harness the ubiquitin proteasome pathway, it is clear that we must be aware of the possibility for crosstalk amongst ubiquitin signaling pathways. With this natural complexity also comes the opportunity to utilize not only the proteasomal degradation pathway, but also alternative pathways of degradation mediated by ubiquitination. In this work, we have focused on harnessing the ubiquitin proteasome pathway by means of engineering the E3 ubiquitin ligase CHIP. In the future, by considering all the natural degradation pathways we may be able to develop a suite of tools for ubiquitination engineering (i.e. proteasomal and autophagic) which could have more widespread applicability. Notably, by working with the ubiquitin ligase CHIP, which naturally interacts with various degradation pathways, we are uniquely poised to diversify our targeted substrate degradation.

Endogenous CHIP is known to interact with heat shock proteins in a concerted effort to prevent misfolded protein aggregation [76]. In a very elegant way, CHIP is able to aid Hsp70 substrates in refolding or target them for degradation by the ubiquitin proteasome pathway, then CHIP can ubiquitinate the co-chaperone Hsp70 to return the cell to an “unstressed” state [67]. However, it has also been shown that CHIP/Hsp70 substrates can be targeted to juxtanuclear aggresomes, considered a “second line defense”, when the intracellular degradation capacity is

exceeded [90]. Interestingly, substrates can be directed to aggresomes by ubiquitin dependent processes (utilizing HDAC6 transport or p62 recognition [58]) or by a ubiquitin independent process (utilizing Hsp70/BAG3 transport [91]). Juxtanuclear aggresomes typically then become autophagosomes and sequentially autolysosomes whereby the substrates are degraded. However the precision of this process is not fully known, as it appears that protein content within the aggresome may affect the efficiency of autophagy and degradation [92]. Notably, CHIP interfaces with both ubiquitin dependent and ubiquitin independent targeting to aggresomes and thus ubiquibodies could harness this robust interface for multifaceted targeted protein degradation (**Fig. 5.1**).

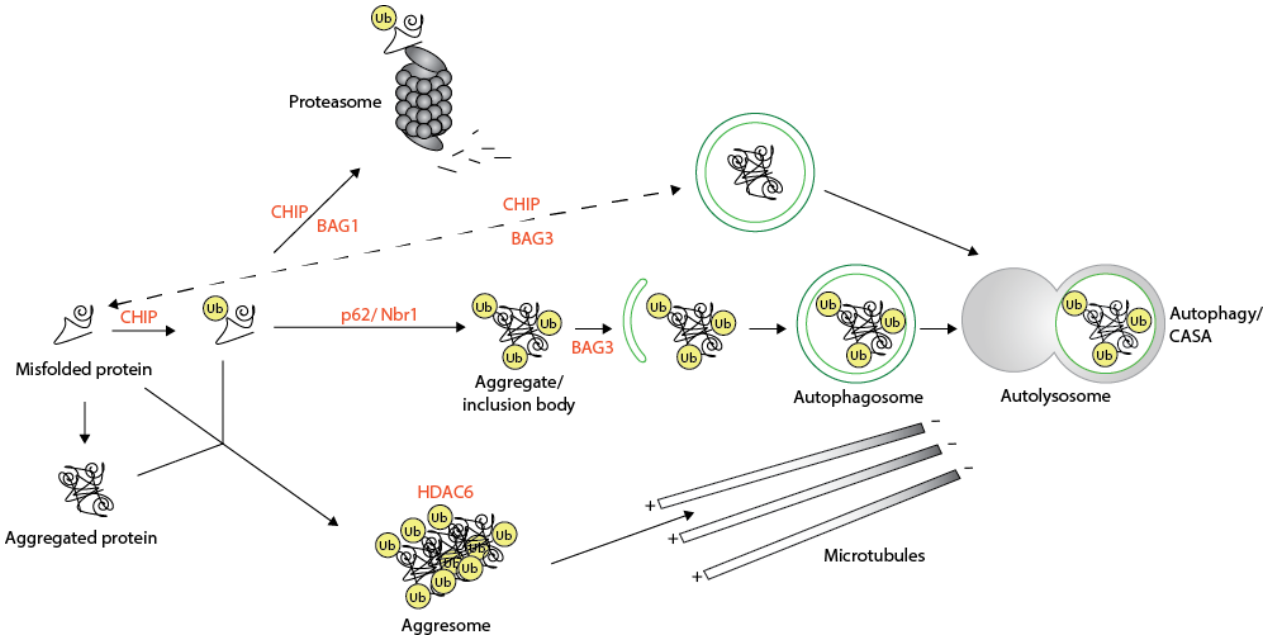


Figure 5.1 Endogenous protein degradation routes mediated by CHIP. As illustrated, CHIP naturally aids in the balance of protein folding, aggregation, ubiquitination and degradation in concert with heat shock proteins and co-chaperones such as BAG. Additionally, misfolded proteins can be degraded by the proteasome or autophagy, in ubiquitin dependent or independent pathways. Illustration was adapted from Kraft et al. [58]**.**

Discussion

One major unanswered question is how are ubiquibodies interfacing with CHIP's natural

signaling pathways? While we removed the TPR domain from CHIP to redirect ubiquibodies, the CHIP homodimer crystal structure reveals that the U-box and helical linker are the dominant dimerization interactions. Thus it is feasible that when uAbs are put into the eukaryotic cellular context they are also forming heterodimers with endogenous wild-type CHIP. As such, this would enable uAbs to tap into CHIP's diverse means of targeting substrates for degradation in concert with heat shock proteins. In order to further elucidate how ubiquibodies are functioning in mammalian cells, a pull-down assay could be used to determine interacting proteins such as wild-type CHIP, Hsp70, and the various E2s.

Possibilities in ubiquibody targeting for autophagic degradation. In light of CHIP's role in the creation of juxtanuclear aggresomes, it is feasible that the detergent-resistant sequestration of β-gal by R4-uAb (see **Fig. 2.11**) was preliminary evidence of ubiquibodies also functioning within the autophagy pathway. Additionally, less β-gal was seen in the insoluble fraction 10 h post-transfection than at 24 h post-transfection using an intermediate level of β-gal transfection (0.1 μg pcDNA3-b-gal, data not shown), which could corroborate the sequential targeting of substrates to the proteasome before autophagic degradation [90]. Notably, we never detected the presence of MBP in the detergent resistant fraction. This could reveal the natural preference for ubiquibodies to target substrate proteins to the proteasome for degradation. However, β-galactosidase forms a tetramer which could, at high expression levels, mimic an aggregation prone protein or overwhelm the proteasomal degradation pathway, thus leading to the second line of defense, autophagy degradation. While we have suggestive evidence towards ubiquibodies functioning within the autophagic pathway, to further study this possibility, immunofluorescent co-localization studies could be performed to label with antibodies against LC3 an autophagosome marker, or vimentin, which shows the intermediate filament cage around

aggresomes [91]. Additionally, it has been shown that the dominant negative point mutant $CHIP^{H260Q}$, primarily sends substrates to the aggresome for degradation [90]. This concept could be expanded upon to create dominant negative ubiquibodies that target substrates resistant to proteasomal degradation due to aggregation propensity [93] or lack of an unstructured handle [94] towards autophagosomes. However, targeting substrates to aggresomes must be cautiously undertaken because cytotoxicity would be a possible undesirable byproduct.

Investigating alternative E3s for targeted degradation. While it could prove powerful to tap into the concerted effort of proteasomal and autophagic degradation pathways, it may also become unnecessarily complicated. An alternative approach would be to utilize E3s with more specialized functions compared to CHIP's natural flexibility in concert with chaperones. One group of E3s to explore would be the HECT (Homologous to E6AP C-terminus) domain containing E3s which contain their own catalytic cysteine which is charged with ubiquitin prior to substrate ubiquitination [16]. The direct transfer of ubiquitin to substrate proteins by HECT E3s suggests they more directly control the polyubiquitin chain linkages built on substrates. Thus by engineering a HECT E3, we could more directly control the substrate polyubiquitination and degradation. To date, the HECT domain of E6AP has been used for ubiquitin ligase engineering [7, 95]. Unfortunately the HECT domain alone is about 39 kDa and HECT E3s range from 100 kDa to greater than 500 kDa in size [96], which may make them more difficult for engineering and analysis by *in vitro* assays. However, the crystal structures of various HECT domains are available and recent studies have revealed a possible mechanism for HECT substrate ubiquitination, which would aid in ubiquibody design [97-99].

Another alternative would be to utilize bacterial E3s with natural specificity for targeting substrates for proteasomal degradation. While prokaryotes do not utilize ubiquitin signaling

pathways as eukaryotes do, pathogenic bacteria have evolved to hijack host ubiquitination machinery for their own benefit. These bacteria deliver effector proteins, also known as virulence factors, into host cells in order to alter the physiological response of the host [100]. A number of these effector proteins have been identified as E3 ubiquitin ligases, with similarities to RING or HECT type E3s. Additionally, a new class of ubiquitin ligases has been identified within bacterial effector proteins, termed NEL for Novel E3 Ligase [101]. Remarkably, these NEL E3s appear to use N-terminal leucine rich repeat (LRR) domains for target binding and have novel C-terminal ubiquitin ligase activity domains with a conserved active site cysteine [102]. Thus, NEL E3s could be a natural scaffold for utilizing LRR designer binding proteins [42] in targeting diverse substrates for proteolytic degradation.

Controlling ubiquibody activity in space and time. Another future direction for ubiquibodies, would be the development of allosteric ubiquibodies with more tightly controlled activity. Specifically, coupling ubiquibody activity to a small molecule ligand or light inducibility would enable temporal or spatial ubiquibody control. Precedence for small molecule induced E3s is present in nature [103] and in previous work studying the flexibility between CHIP's TPR and U-box domains [48]. Additionally, plants utilize light reactive E3s to regulate their circadian clock and photoperiodic flowering [104]. Such allosteric ubiquibodies could promote faster degradation because their response would only require changes in proteins already present in the cell and would not be slowed by having to wait for transcription/translation. As such, these ubiquibody "switches" could be very powerful in synthetic biology networks.

Targeting post-translational modifications. Finally, because ubiquibodies operate at the post-translational level, the ubiquibody technique has the potential for depleting certain protein isoforms while sparing others. Supporting this idea is the finding that natural E3s recognize

specifically phosphorylated proteins in cell cycle regulation, thus coupling proteolysis to cyclin-dependent kinases [26]. Additionally, *N*-linked high-mannose oligosaccharides are an endogenous signal for endoplasmic reticulum associated degradation E3s [105]. In fact, many post-translational modifications (PTMs) regulate substrate recognition by E3s, including hydroxylation, deacetylation and oxidation [106]. Amazingly, the natural diversity of E3 ubiquitin ligases has reminded us that "if you can think of it, nature has already tried it a long time ago" [42]. Thus it should be natural to extend ubiquibody application towards studying post-translationally modified proteins. Specifically, DBP domains which can selectively bind isoforms or PTMs could be employed to degrade and further elucidate distinct functions [107, 108]. This is significant because isoform targeting is difficult, and PTM-specific silencing is not possible using current knockdown methods that function at the level of DNA or RNA.

Conclusions

In this work, we have presented a generalizable approach for protein knockdown by developing ubiquibodies, which combine the activity of an E3 ubiquitin ligase with the specificity of designer binding proteins. While the ubiquibodies developed to date show the potential diversity and robustness of targeted protein degradation, they also reveal potential areas for improvement (e.g. binding kinetics and E2 specificity). However future endeavors may be far more ambitious than continuing to optimize CHIP based ubiquibody design and could venture into alternative E3 scaffolds, allosteric ubiquibodies or targeting post-translational modifications. With such endeavors, the utility of ubiquibodies could be expanded towards studying natural protein signaling pathways dependent on temporal or post-translational regulations. As such, the ubiquibody technology has the potential to enable the further dissection of protein functions or be used to selectively degrade proteins that underlie human disease.

Acknowledgements

I'd like to thank the many teachers who have taught me the foundational science underlying all of this research, especially Professors Linda Nicholson, Bill Brown and Susan Ely at Cornell and Professors Phil Campbell and Gordon Rule at Carnegie Mellon. Additionally many of these ideas have come from valuable conversations with Matt DeLisa, Sean O'Brien, Jason Boock, Erin Stephens and Morgan Baltz. This material is based upon work supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Institutes of Health Chemical Biology Training Grant T32 GM008500 (fellowship to A.D.P.) and the National Science Foundation GK-12 DGE-0841291 (fellowship to A.D.P.).

CHAPTER 6

GK12: FORENSIC DNA FINGERPRINTING

Introduction

Cornell's Learning Initiative in Medicine and Bioengineering (CLIMB) is an NSF funded graduate program for kindergarten through 12th grade education (GK-12) based out of the Department of Biomedical Engineering. The goal of CLIMB is for graduate students to partner with a middle or high school teacher and develop inquiry based curriculum to engage students in science, technology, engineering and math (STEM). For my GK-12 experience, I was partnered with David Syracuse of Groton High School. Over a six-week summer program, Dave and I participated in GK-12 courses focused on inquiry based learning and curriculum development. During this time, Dave also joined me in the DeLisa research labs to learn about and participate in my thesis research. While engaging in research together, we sought to identify areas where my research and his classroom curriculum would intersect. During the 2009-2010 academic year, Dave taught Living Environment, NY state's required basic biology course, in addition to two senior level courses, Forensic Science and Environmental Science. I had weekly interactions in each of these classrooms, implemented a citizen science program with the Environmental Science course, and designed and implemented a Forensic DNA Fingerprinting curriculum.

Weekly interactions as the "scientist in residence"

Throughout the academic year I visited Dave's classes on a weekly basis, either expanding on the subject matter with mini-lectures or aiding in classroom laboratory experiments, depending on the curriculum. These interactions were critical for me to learn the level of scientific understanding amongst the students and for the students to engage with me as a scientist. When I presented material to Dave's classes, I tried to integrate my research with their

course material. For example, I explained how we quantify proteins in the lab as they were learning to detect the presence of proteins. I also tried to expand upon Dave's standard teaching techniques, such as showing them a 3D crystal structure of salivary amylase when they were conducting a lab on this enzyme. The students really connected with this interactive medium because they could rotate and move the crystal structure model. We also related this to their study of protein secondary structures which was typically taught using 2D drawings on paper.

I not only expanded upon their classroom material, but also asked them to engage in my experimental coursework, in an effort to show them ongoing learning. For example, while they were learning about light microscopes, I was learning to use a two-photon microscope for a class. So I taught them in general terms what two-photon microscopy was, and then asked them to suggest what I should look at under the microscope. The class decided it should be a bug with wings. I then brought in pictures of the microscope set-up , images of the bug and shared my experiences with them of both the excitement and disappointment in the images. This really demonstrated how much easier it was to get students interested in topics that I also had a vested interest in, whether due to my research or coursework.

Wolbachia curriculum in Environmental Science

The *Wolbachia* Project called "Discover the Microbes Within!" is a nationwide program sponsored by the Marine Biological Laboratory designed to engage high school students in real-world scientific research. The research aim of the program is to analyze local insects for the microbe *Wolbachia pipientis* in order to track infected species. The project coordinates workshops across the country to train teachers in the curricula and Cornell's Institute for Biology Teachers (CIBT) also supports this program. Dave and I took the curriculum training through CIBT and brought this unique program to his Environmental Science course during the spring

semester.

At the beginning of the *Wolbachia* program, we shared with the students how tracking species infected with the bacterial endosymbiont *Wolbachia pipientis* relates to human health. Currently, *W. pipientis* is making headline news as a potential "mosquito vaccine" towards reducing the transmission of dengue fever which is an infectious tropical disease. *W. pipientis* is considered a "reproductive parasite" because it alters sexual reproduction, but it is also able to make mosquitoes immune to the dengue virus and is passed on to progeny. Thus by releasing a population of infected mosquitoes, the local population of mosquitoes carrying dengue can be significantly reduced [109]. Additionally, *W. pipientis* is known to infect nematodes which can cause heartworm in dogs and river-blindness in human beings. These real-world implications of *W. pipientis* were used to excite the students for the program and to give them an idea of how seemingly obscure research can contribute to health and medicine.

The *Wolbachia* project provided the students an opportunity to collect insects and classify them using a web based identification key. They then extracted DNA from the insects and performed a polymerase chain reaction (PCR) to amplify the *Wolbachia pipientis* DNA. Finally they evaluated their PCR products using gel electrophoresis. This unit was specifically geared toward the Environmental Science curriculum and was a unique opportunity for the students to engage in a citizen science project.

Forensic Science DNA Fingerprinting curriculum

This unit was organized over five class periods with two introductory lessons, two experimental lessons and one conclusion lesson. The overall goal of the unit was to reinforce genetic concepts by allowing the students to perform DNA-based molecular biology experiments to determine which suspect was at the scene of a crime. After learning about DNA and gel

electrophoresis, the students were able performing polymerase chain reaction (PCR) to amplify the DNA collected at the "crime scene". In Forensic work, PCR amplification of DNA allows us to look at loci with alleles that differentiate people. These loci, known as short tandem repeats (STRs), are also used by the FBI to identify criminals. We also discussed the power of discrimination, highlighting that looking at up to 13 different loci is required to identify one individual from millions of people. The curriculum and handouts used for this program are included as Appendices A and B, respectively.

During this curriculum I noticed how the students who had previously been disengaged became truly intrigued in the new material and experimental lab set-up. I was surprised to find that the students, who had historically struggled to grasp concepts, became more engaged, as evidence by their willingness to ask questions. The students were also very attentive and wanted to ensure they were doing all the hands-on techniques correctly. Although the results from the experiments were a little unclear (literally fuzzy pictures) the students completed their worksheets for the unit and asked questions regarding topics they were unsure about. Overall it was a great experience and very rewarding to see the students engage in a technical topic full of scientific techniques used in both molecular biology and forensic analyses.

Student Evaluation

In an attempt to analyze how the DNA Fingerprinting unit influenced the students, I utilized a survey before and after the unit that included both a test of student attitudes and content in the genre of genetics (Appendix C). Questions were taken directly from two standardized tests: the TOSRA (Test of Science-Related Attitudes) subsets on Social Implications of Science, Adoption of Scientific Attitudes and Attitude to Scientific Inquiry and a subset of questions from Dr. Susan Elrod's Genetics Concepts Inventory (GCI) [110, 111]. I selected questions from the GCI based

on the material I knew we would touch upon in the unit. However, in comparing the before and after surveys, no significant improvements were made in the content of genetic concepts. In retrospect, I realized while the questions drew on fundamental genetic concepts, the material was not specifically covered during my unit and therefore would require the students to refer back to previous biology courses. In comparing the TOSRA subsets before and after, a slight increase in attitudes towards science was seen in all areas, as shown below. The data was averaged over seventeen students who took both the pre- and post-surveys (with parental permission). Questions asked in a negative way were corrected for as suggested by the TOSRA handbook [111]. While the improvement of attitudes seen is not statistically significant, it could reveal the overall influence of a positive experience on the students' inclinations toward science.

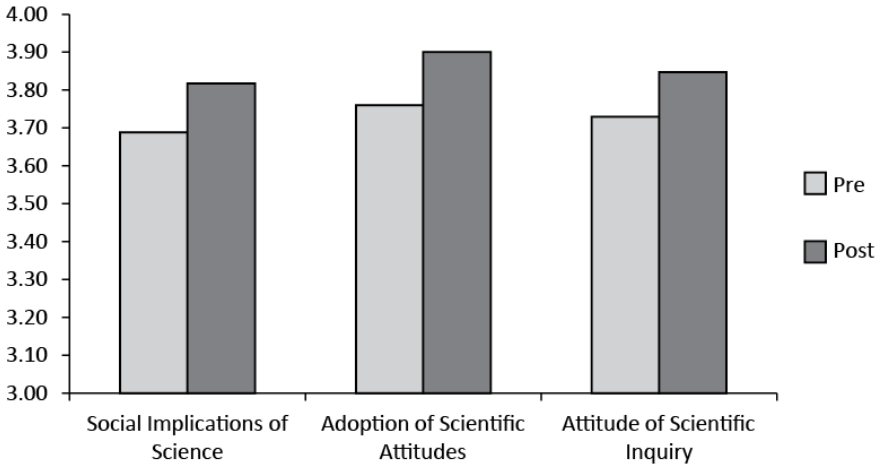


Figure 6.1 Test of Science-Related Attitudes. Seventeen students from the Forensic Science course participated in pre- and post-curriculum surveys evaluating their attitudes related to science. The scores from these surveys were clustered into topics including the social implications of science, adoption of scientific attitudes and attitude of scientific inquiry. In these three subsets, a marginal improvement of attitude was detected.

Discussion

The goal of the National Science Foundation in supporting graduate students through GK-12 fellowships is to two-fold. First, graduate students are given a unique opportunity to teach

science and communicate at an educational level unlike standard training opportunities such as university teaching assistance appointments. Secondly, by engaging in K-12 education, graduate students have an opportunity to encourage young burgeoning students in STEM content. The entire program is built to be a two-way street, where K-12 students get excited about science and graduate students learn to teach and communicate. In my GK-12 experience, I benefited greatly by the unobserved component of this system, my high school teacher Dave Syracuse. Dave is a very enthusiastic and engaging science teacher who daily tried to excite his students in scientific materials. As such, I believe I benefited far more from this GK-12 exchange.

Acknowledgements

This GK-12 experience would not have been possible without the support, encouragement and direction of Dave Syracuse. Additionally thanks are due to the participating high school students at Groton High School who probably taught me more about the range of teaching than I taught them about science. Many thanks to Professors Chris Schaffer, Mike Shuler and Shivaun Archer who developed the BME GK-12 program, taught courses on inquiry based learning and guided curriculum development. Also thanks are due to Nev Singhota who coordinated the summer program, speakers and the acquisition of teaching materials. Additionally this experience was greatly enhanced by fellow 2009-2010 graduate students and teachers working to challenge each other and engage students. Finally I'd like to acknowledge Dr. Seth Bordenstein for developing the *Wolbachia* curriculum and BioRad for the framework of the curriculum used in the DNA Fingerprinting unit. This material is based upon work supported by the National Science Foundation GK-12 DGE-0841291 awarded to Professor Shuler of the Biomedical Engineering Department at Cornell (fellowship to A.D.P.).

Forensic DNA Fingerprinting

Subject: Forensic Science/Advanced Biology
Level: High School
Standards: *New York State- Living Environment*
Standard 1- Analysis, Inquiry and Design
Standard 4.2- Genetic Inheritance, 4.3- Organisms change over time, 4.4- Continuity of Life
Schedule: 5 days (1 content, 3 experimental & 1 wrap-up)

Objectives:

Students will understand the process of synthetic DNA replication by polymerase chain reaction. They will also understand agarose gel electrophoresis and the capability to separate DNA according to its size.

Students will:

- Use their understanding of electrophoresis to determine the charge of dyes based on their separation. And to determine the dyes in a mixed unknown.
- Implement the process of Polymerase Chain Reaction by mixing the necessary components and running the reaction on a thermal cycler.
- Analyze their PCR samples by running them on an agarose gel electrophoresis to determine which suspect was at the crime of the scene.

Vocabulary:

Polymerase Chain Reaction (PCR)
DNA Gel Electrophoresis

Materials:

- Student handouts (daily)
- 20uL adjustable pipettes P20 (student use)
- 200uL adjustable pipette P200 (instructor use)
- Gel electrophoresis tanks
- Power supply
- Biorad Crime Scene Investigator PCR Basics™ Kit (#166-2600EDU)
- Marking pens (sharpies)
- Distilled water
- Water bath
- Microcentrifuge
- Thermal cycler
- Biorad Small Electrophoresis Reagent Pack (#166-0450EDU)
- Food colored 10% glycerol solutions (optional day 2)
- Dyes for migration study (optional day 2)- available from Frey's (95-0420-016)
- Eppendorfs (optional day 2)
- SybrSafe & Blue Light Box (optional alternative FastBlast blue stain from kit)

(See Classroom Procedure for daily lists)

Safety:

Gloves should be worn at all times in order to avoid contaminating samples and to prevent contact with bases and EDTA (in gel & buffer).

Day 1: Introduction to PCR and Electrophoresis (42 minute lesson)

Objectives (Content/Process): Students will be able to explain in their own words principles behind PCR and electrophoresis. They will also be able to plan an electrophoresis experiment, and why it would be pertinent in a scientific setting.

Materials:

- Handouts about DNA structure, PCR and electrophoresis
- DNA replication video, available on Youtube: http://www.youtube.com/watch?v=teV62zrm2P0&feature=related)
- Nickelodeon guts video, available on Youtube: http://www.youtube.com/watch?v=NcxjUq6m8Tg&feature=related
    - Clips at 5:40 & 6:24 minutes showing "elastic jungle" which gives a visual parallel to DNA traveling through an agarose gel (larger is slower)

Lesson:

1. What has to happen for DNA replication to occur?
    a. Open the two strands to get access to the bases ATC&G – **topoisomerase** unwinds the DNA and **helicase** separates the two strands
    b. **Single strand binding proteins** line up along the DNA to hold the two strands apart
    c. Now we need to add new bases to the separated strands, this is initiated by a **RNA primer** which is added by an enzyme called **primase,** the RNA is eventually replaced by DNA though
    d. DNA extension can then continue after the primer and is done by **DNA polymerase**, one by one nucleotides are aligned and added at the 3' end… what about the other side of the fork?
    e. The two strands are called leading and lagging strands, the leading strand has nucleotides added directly, the lagging strand has what are called **Okazaki fragments** which are then "glued" together by **DNA ligase**
    f. Show YouTube video
2. So why is replication important to forensics? We need lots of DNA to identify an individual, but often there are trace amounts at the scene of a crime
3. Question to the students: What can scientists do with DNA?
    a. Other questions… why do they want to work with DNA? What is the significance?
    b. Cloning – make synthetic DNA or move DNA from one organism to another
    c. Genetic engineering – alter genes, therefore proteins, therefore functioning
    d. Therapeutic uses – possible therapeutic uses include gene therapy or biological therapeutics such as proteins (possibly engineered genetically)
4. Techniques for working with DNA – used by both research and forensic scientists
    a. PCR: polymerase chain reaction
        i. Allows amplification of a specific region on DNA – why would it be valuable to copy DNA?

1. 1 copy of DNA -> millions of copies
2. Increased number of copies exponentially
3. Cheap and easy technique, therefore revolutionized molecular biology

ii. Required Materials:
1. Taq polymerase - from *Thermus aquaticus*a thermophilic bacteria from which it was isolated in 1965; optimum temperature for activity is 75-80˚C
2. dNTPs – deoxyribonucleotide triphosphates
3. buffer (w/Mg required cofactor of polymerase)
4. primers – 18-22bp
5. template DNA

iii. 3 steps of PCR:
1. Denaturation – 2 strands separate at 90-95C
2. Annealing – primers anneal to template strands at 55-70C (depends on the primer length and composition, Tm indicates temp at with ½ is duplex and ½ is single stranded)
3. Extension – polymerase adds A, T, C, G at 72C

b. Electrophoresis
i. Gain interest by showing video of Nickelodeon Guts bungee obstacle course
ii. Motion of particles relative to a fluid under the influence of a uniform electric field
iii. Movement is cause by the presence of a charged interface between the particle surface and the surrounding fluid
iv. In a negative to positive field, which way would DNA travel? To the positive end
v. By adding a porous medium, agarose gel, the size of the DNA now significantly affects the migration of the negatively charged DNA
vi. Small pieces travel faster than large pieces of DNA
vii. This allows the separation of different sized pieces of DNA
viii. Main principles of agarose gel electrophoresis
1. Negative charge, size of fragment, concentration of gel
2. Voltage – how quickly do the charged particles move
3. Buffer – conductive material

Check for comprehension: Ask students to draw an electrophoresis setup

---

Day 2: Techniques of Electrophoresis (42 minute lesson)

Objectives (Content/Process): Students will be able to setup a DNA gel electrophoresis and properly handle all required equipment including a gel rig, pipette, samples, and voltage source.

Materials:

- P200 to aliquot samples (teacher)
- 20 P20s for students
- Pipette tips & boxes
- Parafilm
- Dyed glycerol solution
- Colored tape
- Precast 1% agarose gels (one per group) with combs in the **middle** (200mL gel @ 1%)
- Dyes for migration:
    - solution A & B (mixture of dyes as unknowns)
    - xylene cyanol (538.6g/mol), negative ~4kb in 1%
    - orange g (452.4g/mol), negative ~ 50bp
    - methyl green (458.5g/mol), positive
    - bromophenol blue (669.96g/mol), positive ~500bp

Experimental Lab:

1. Pipetting techniques
    a. First determine the volume you want to pipette and set the dial
    b. Pipettes have two stops, the first is the correct volume, the second adds a burst of air to eject any remaining solution
    c. Go to the first stop when you push air out before pipetting up
    d. Pipette up **slowly** to ensure that no liquid gets up into the pipette – this can ruin the seal
    e. When pipetting out, use the second stop to get the last drop out
    f. Pipette colored water drops of varying volumes onto parafilm (5, 10)
    g. Mixing: pipette slowly and only go to first stop
    h. Mix your two favorite colors
    i. Check the bottom of your pipette to make sure no color is on the bottom, if there is, you've pipetted too quickly or too much!
2. Running a gel with dyes to see migration
    a. One precast gel per group
    b. Set up the gel in the rig and remove the comb
    c. Dilute buffer from 10x to 1x and pour into gel tank
    d. Load each well with 10uL of dye, using a separate tip for each well, making sure to write down the order of your samples being loaded
    e. Run the gel at 200V for quick migration (~15min)
    f. Evaluate dye migration of controls compared to the unknown samples

Check for comprehension: What were the dyes in the unknown solutions? What do you expect to see from a DNA fingerprinting gel? What will be the same as this dye gel and what will be different?

Day 3: Polymerase Chain Reaction (42 minute lesson)
Objectives (Content/Process): Students will be able to explain why PCR is used in science and how it works. Also they will know all the ingredients required for a PCR to work and the three steps of the reaction.

Materials:

- P200 & P20 for teacher to aliquot samples
- P20s for students
- Tubes – with DNA aliquots and MasterMix + Primer (MMP) premixed
- Tube racks
- Small tube racks
- Pipette tips
- Tube markers - sharpies
- Ice boxes
- Thermal cycler programmed
- Handout for students with protocol and questions

Experimental Lab:

1. Review Theory:
    a. Required Materials:
        i. Taq polymerase - from *Thermus aquaticus*a thermophilic bacteria from which it was isolated in 1965; optimum temperature for activity is 75-80˚C
        ii. dNTPs – deoxyribonucleotide triphosphates
        iii. buffer (w/Mg required cofactor of polymerase)
        iv. primers – 18-22bp
        v. template DNA
    b. 3 steps of PCR:
        i. Denaturation – 2 strands separate at 90-95C
        ii. Annealing – primers anneal to template strands at 55-70C (depends on the primer length and composition, Tm indicates temp at with ½ is duplex and ½ is single stranded)
        iii. Extension – polymerase adds A, T, C, G at 72C
2. Label all tubes – MMP, CS, A, B, C, D + group identifier
3. Keep all samples cool on ice
4. Pipette 20uL of each DNA sample into a PCR tube, be sure to use a fresh pipette tip for each

5. Add 20uL of MMP to each tube, using a fresh pipette tip each time
6. Close the tube and mix by flicking the tube
7. Start thermal cycler
    a. Review 3 steps and the program of the thermal cycler

Lesson:

1. Allele Frequencies & STR loci
    a. Alleles are alternate forms of a gene, original word was allelomorph meaning "other form"
    b. A gene is a discrete unit of hereditary information consisting of a specific nucleotide sequence in DNA
    c. A gene may produce a product of RNA or protein … somewhat complex
    d. An allele describes the variance among people, we all have similar genes which perform the same function, but the sequence may be slightly different
    e. By looking at specific loci (locations in the genetic material) and amplifying that gene, alleles can be detected
    f. For example, blood typing is base on the different alleles, AB&O
        i. The ABO locus is located on chromosome 9 and has three allelic forms encoding two different forms of a protein called glycosyltransferase that bonds different sugars to the H antigen in red blood cells, the O allele has a mutation such that no glycosyltransferase is produced and the H antigen remains unmodified
    g. Some loci contain very recognizable and variable sequences called short tandem repeats or STRs
    h. STRs are short repeats of the same sequence : ATCA ATCA ATCA ATCA
    i. What is recognizable is that the number of repeats varies from person to person and often within each person on their chromosomes
    j. Draw what this looks like on the board
    k. In forensic analysis often times 13 different loci will be analyzed to increase the power of discrimination
    l. In our PCR we have used primers that bind one loci, BXP007 and will show us the alleles of DNA from the crime scene and from suspects, there are 8 different possible alleles (15, 10, 7, 5, 4, 3, 2, 1)
    m. STRs can also be used to determine familial relationships

Check for comprehension: handout with questions regarding alleles and PCR

Day 4: Gel Electrophoresis & Analysis (42 minute lesson)

Objectives (Content/Process): Students will be able to setup a DNA agarose gel electrophoresis and properly handle all material required.

Materials:

- 3 x 250mL, 3% agarose gels- pre-cast with SYBR Safe
- Flask to make TAE solution in
- Gel rigs
- Power sources
- Student P20 pipettes
- Pipette tips
- PCR samples
- Orange G loading dye
- Allele Ladder
- 0.5x TAE (from 50x stock)
- SYBR Safe blue light boxes
- Camera
- Handout for students with protocol and questions

Experimental Lab:

1. Centrifuge PCR tubes to get all solution in the bottom
2. Add loading dye 10uL to each PCR mix
3. Check orientation of gel in the gel tank
4. Order samples for loading and write down the order
5. Load samples individually – 20uL each
6. Secure lid in correct orientation
7. Turn on power supply and electrophorese gel at 200V for 20 min
8. Look at gel after electrophoresis with blue light box and orange shield
9. Take picture of gel
10. Document each suspect's allele pattern
11. Identify suspect who was at the crime scene

Check for comprehension: handout with questions regarding gel electrophoresis

---

Day 5: Wrap-up/ Conclusions (42 minute lesson)

Objectives (Content/Process): Ensure students understand the process of DNA Fingerprinting and the complexity of using this technique to accurately identify a culprit from a pool of suspects.

Materials:

- Pictures of gels
- TH01 allele frequency image – from BioRad handbook p.10
- Loci visual aid of power of discrimination
-

Lesson:

- Power of Discrimination:
    - Larger the number of loci typed, the more powerful the ability to discriminate between individuals
    - Allele frequencies don't follow regular mathematical patterns
    - Allele frequencies change among different populations – look at TH01 allele frequency graph
    - Can use the frequency of each allele to determine the frequency of a genotype
    - Example: Looking at multiple loci
        - A Caucasian suspect has STR of 16 & 17 at loci D3S1358, the frequencies of which are 16=0.253 and 17=0.215 in the Caucasian population
        - The frequency of this 16-17 genotype can be calculated by multiplying :
          2 x 0.253 x 0.215 = 0.109 or 1/10 people
        - 1/10 people is not very convincing … let's look at a second loci
        - Same suspect has a STR of 10-8 at loci TH01, the frequencies of which are 10 = 0.008 and 8 = 0.125
        - The frequency of this 10-8 TH01 genotype is :
          2 x 0.008 x 0.125 = 0.002 or 1/500 people
        - 1/500 people is more convincing, but if we use both the D3S1358 and the TH01 locus (0.125x0.002 = 0.000218) we can say it's found only in 1 out of 5,000 people
        - Note that the FBI uses 13 loci!
- CODIS: Combined DNA Index System – maintained by FBI to help connect cases across states both criminal and missing persons
- Actual complexity of loci: http://www.cstl.nist.gov/biotech/strbase/str_fact.htm
- Mitochondrial DNA matching: for missing persons (inherited only from mother) therefore not unique enough for forensics
- Power of discrimination statistics:

Check for comprehension: post-survey to see concept comprehension regarding genetic information such as genes, alleles and loci

---

Extension Activities:

See the Crime Scene Investigator PCR Basics™ Kit Handbook which contains more materials for lectures (Appendix A) as well as a sample case (Appendix B) and an extension of looking at the power of discrimination (Appendices B & C).

Supplemental Information:

DNA in the electrophoresis gels can be visualized by FastBlast stain (overnight or in 20min) or using SybrSafe in the gel. If using FastBlast stain, the evaluation of the results will not occur until Day 5. Also, day 4 and day 5 should be back to back to avoid diffusion.

Safety:

General Practices: wear gloves at all times when handling buffers and solutions

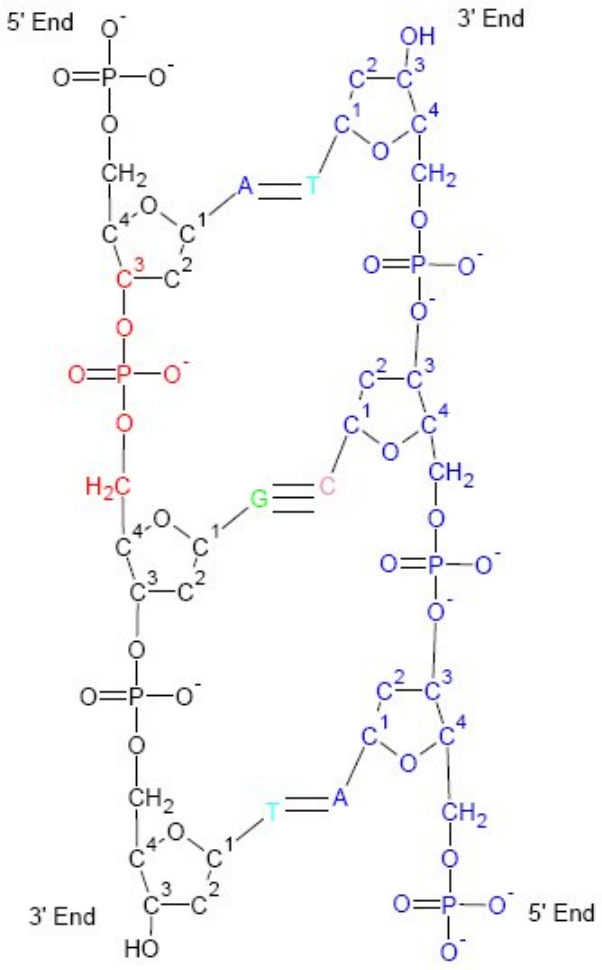
SybrSafe – considered non-carcinogenic, with a toxicity LD50> 5000mg/kg. However it still chelates with DNA and therefore should be handled very carefully, wearing gloves at all times. Disposal should be arranged with an Environmental Health & Safety Office.

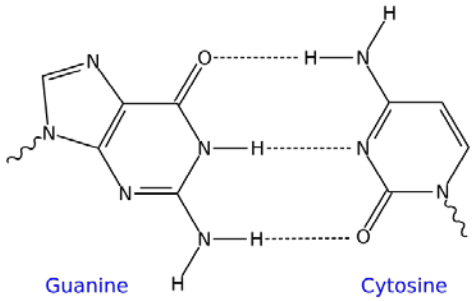
TAE – contains Tris Base, Acetic Acid and EDTA, all of which are skin irritants and thus gloves should be worn at all times.

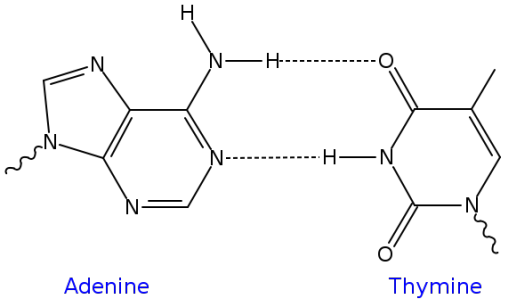
Agarose –skin irritant, flush with water
xylene cyanol – skin irritant, flush with water
orange g – skin irritant, flush with water
methyl green – skin irritant, flush with water
bromophenol blue – skin irritant, flush with water

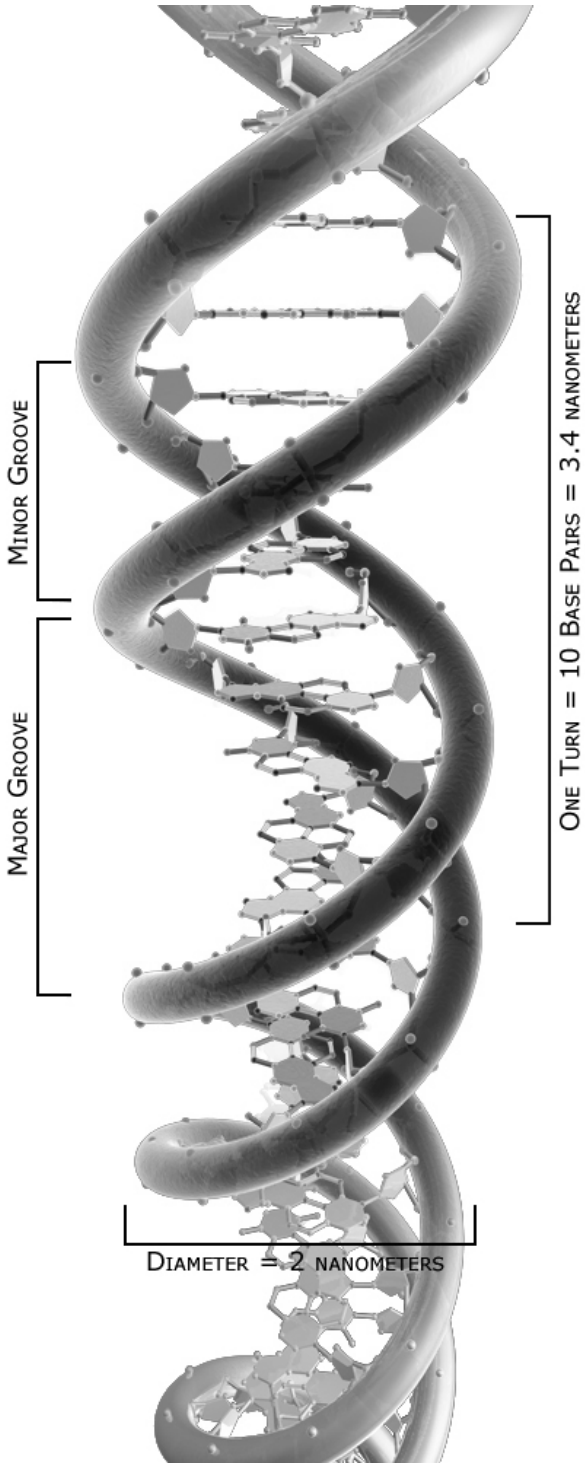
Waste Disposal Method - All chemical disposal must be in accordance with current local, state, and federal regulations.

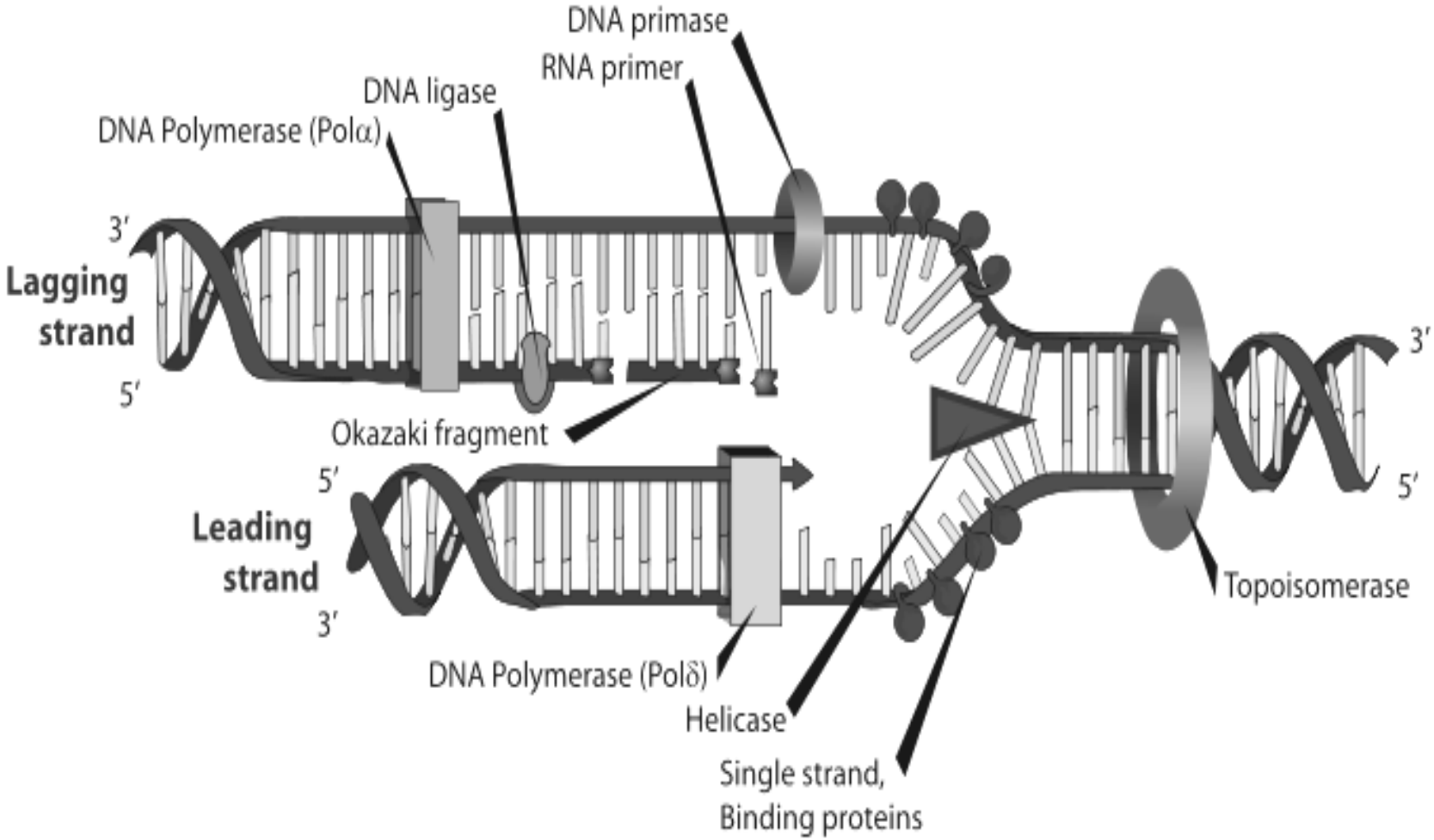
DNA Structure









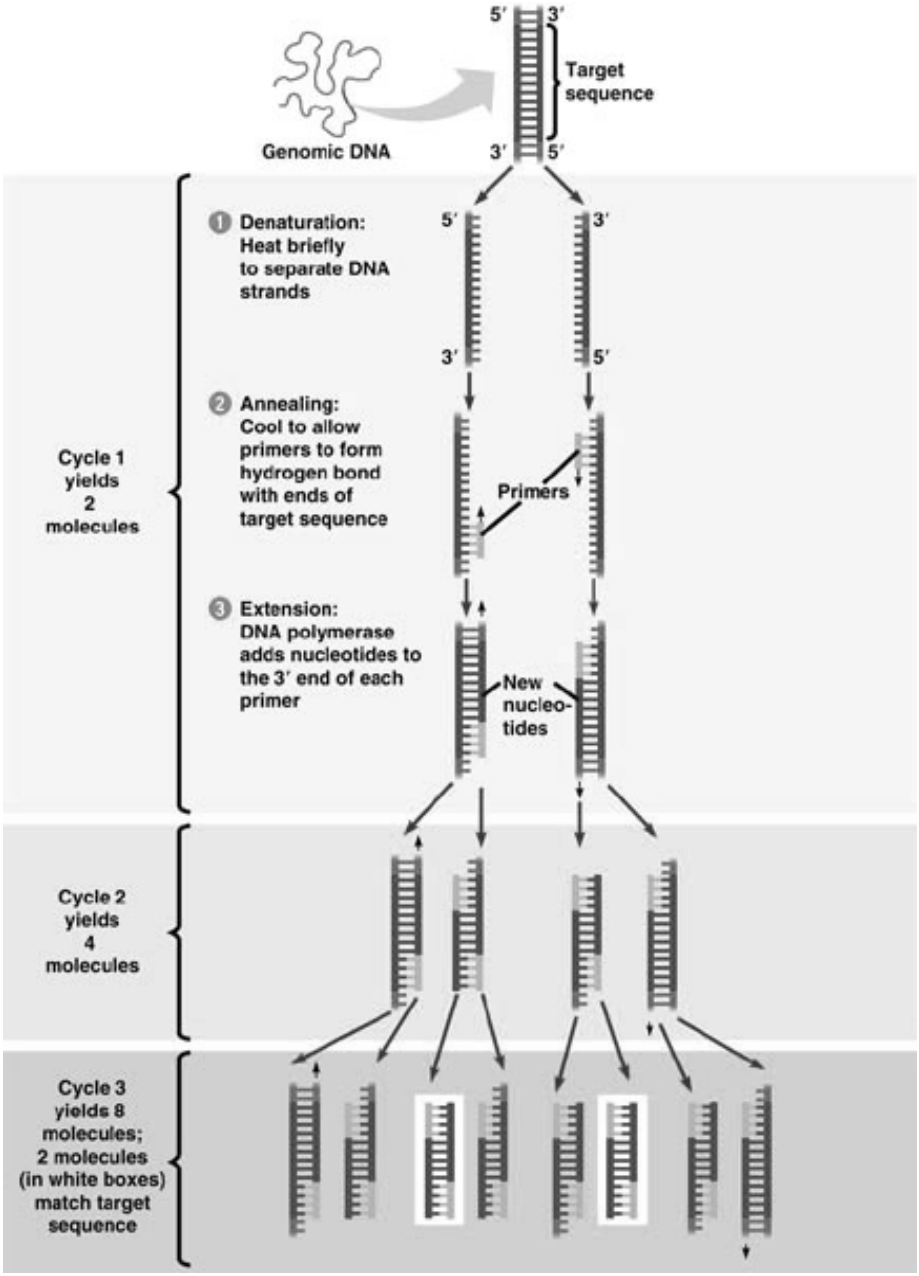

DNA primase
RNA primer
DNA ligase
DNA Polymerase (Polα)
3′
Lagging strand
5′
Okazaki fragment
5′
Leading strand
3′
DNA Polymerase (Polδ)
Helicase
Single strand, Binding proteins
Topoisomerase
3′
5′

DNA Fingerprinting Techniques

Polymerase Chain Reaction (PCR):

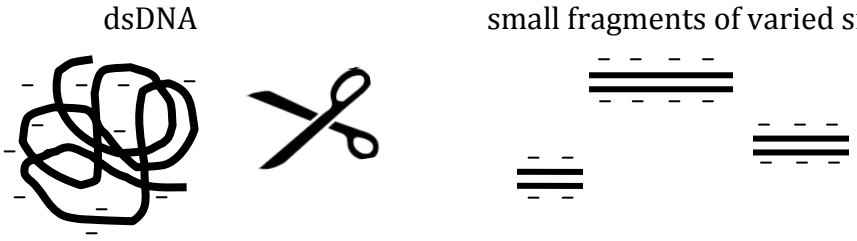
Materials:
1) Taq DNA polymerase
2) dNTPs (deoxyribonucleotide triphosphates) – A, T, C, G
3) primers (18-22bp)
4) template dsDNA (double stranded DNA)
5) buffer solution

Process:

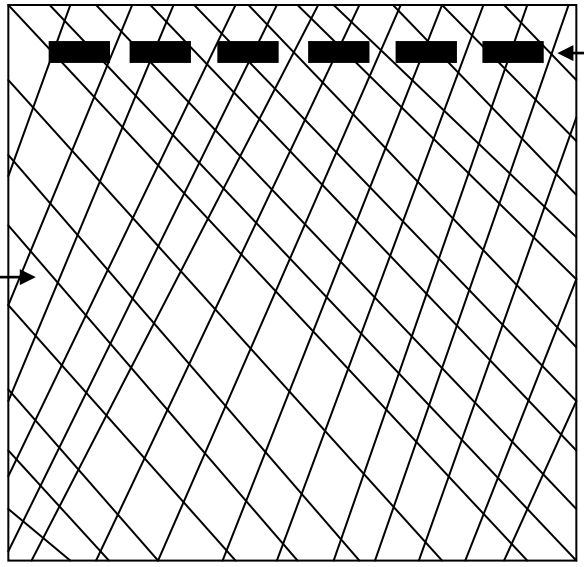


25 rounds -> **2^{25} = 33,554,432 copies of DNA!**

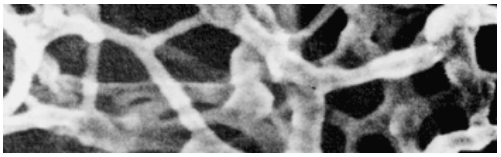
DNA Agarose Gel Electrophoresis:



(-) negative charge (-)



agarose matrix



(+) positive charge (+)

DNA electrophoresis results:

Crime Scene: Highway Motel Rm#13

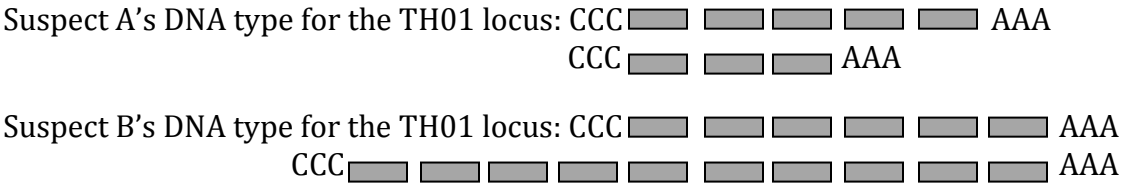
The motel manager hears loud voices, a woman screams, and a shot rings out. The manager runs to the window in time to see the receding lights of a car leaving in a hurry. The door to room #13 hangs open. The manager runs to the open door, to see a man lying face down in a pool of blood. He calls 911. The police arrive, and begin to examine the crime scene. An apparent homicide, but with no obvious clues as to who committed the crime. Or… ?

What kinds of human DNA sequences are used in crime scene investigations? There are ~3billion basepairs in the human genome – greater than 99.5% don't vary between different human beings. However, a small percentage of the human DNA sequence (<0.5%) does differ, and these are the special ***polymorphic*** ("many forms") sequences used in forensic applications. By universal agreement, DNA sequences used for forensic profiling are "anonymous"; that is, they come from regions of our chromosomes (also called ***loci***) that do not control any known traits and have no known functions. Loci are basically genetic addresses or locations. A single ***locus*** may have different forms or types; these different forms are called ***alleles***. A locus may be bi-allelic, having only two different forms, or it may be polymorphic, as described above.

The DNA sequences used in forensic labs are non-coding regions that contain segments of ***Short Tandem Repeats*** or ***STRs***. STRs are very short DNA sequences that are repeated in direct head-to-tail fashion. The example below shows a locus (known as TH01) found on chromosome 11; its specific DNA sequence contains four repeats of [TCAT].

…CCC TCAT TCAT TCAT TCAT TCA…

For the TH01 STR locus, there are many alternate polymorphic alleles that differ from each other by the number of [TCAT] repeats present in the sequence. Although more than 20 different alleles of TH01 have been discovered in people worldwide, each of us still has only two of these, one inherited from our mother and one inherited from our father. For example as shown below, suspect A has one allele with 5 repeats, and one allele with 3 repeats, giving a DNA profile for the TH01 locus of 5-3.



Imagine a scenario in which Suspect A and Suspect B are accused of being involved in a love triangle and committing the murder of a third person in the Highway Motel Room #13; the person who actually pulled the trigger is unknown. In addition to DNA samples from the crime scene, the forensic specialist will isolate DNA from suspects, victims, and others present to genotype as controls. Using PCR-based analysis, the samples will be examined at 13 different genetic loci, using software to interpret the results from the amplification products. In real crime scene analysis DNA profiling is performed at many loci to improve the ***power of discrimination*** of the testing. The power of discrimination is the ability of the profiling to tell the genetic difference between different individuals. The larger the number of loci profiled, the more powerful the ability to discriminate.

You are about to conduct real world forensic DNA profiling. As a crime scene investigator, you will use the polymerase chain reaction (PCR) and agarose gel electrophoresis to analyze the DNA samples obtained from a hypothetical crime scene and four suspects. Your job is to identify the perpetrator. A genotype is the particular set of genetic markers, or alleles, in a DNA sample. Every person's genotype is their own uniquely personal genetic barcode. In this experiment, you'll be revealing the genetic barcodes of several individuals by looking at a single locus BXP007 and looking for whodunit!

Polymerase Chain Reaction:

Materials:

- Ice bath containing tubes
- Master Mix + primers (MMP, blue liquid)
- Tubes of DNA (Crime Scene and Suspect A-D DNAs)
- PCR tubes (small tubes)
- Marking pen
- P20 pipette
- Pipette tips

Protocol:

1) Keep tubes on ice during the entire procedure. Only remove them to remove or add a solution.
2) Label the PCR tubes CS, A, B, C and D and include your group name as well.
3) Set your pipettes to 20µL and transfer 20µL of MMP into each of your tubes.
4) Next transfer 20µL of template DNA into the appropriately labeled tube. For example, transfer 20 µL from the Crime Scene tube to the tube labeled 'CS'.
   <u>Important:</u> use a fresh pipette tip for each DNA sample!
5) The solution in your PCR tubes should be blue now. Be sure to cap the tubs securely and keep them on ice.
6) When instructed to do so, place your tubes in the thermal cycler.
7) Write down the cycle that will be used in the thermal cycler.

Questions:

1) What kind of materials obtained from a crime scene might contain DNA?

2) What might you see if you ran a DNA sample extracted from evidence on a gel before PCR? Why do you need to perform PCR on DNA evidence?

3) What is a genotype?

4) What is the difference between an allele and a locus?

5) Why do forensic labs analyze non-coding DNA and not genes?

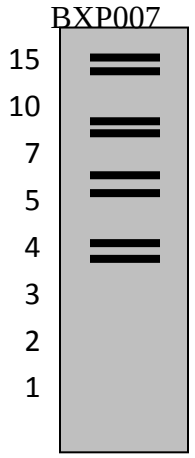
6) What components do you need to perform PCR and why do you need each component?

7) What steps make up a PCR cycle, and what happens at each step?

Crime Scene: DNA evaluation

You have completed your PCR amplification. However, at this point, you can't actually tell whether or not you have PCR products. To do this, you must sort your PCR products using gel electrophoresis and then visualize them using a DNA stain. Since DNA is negatively charged, it can be separated using an electric current. In fact, electrophoresis means "carry with current". In agarose gel electrophoresis, DNA is placed in solidified agarose, which forms sieves containing pores that vary in size depending on the concentration of the agarose. The higher the concentration of agarose, the smaller the pore size, and the longer it takes for larger molecules to move through. This is particularly useful when you want to compare DNA molecules of different sizes contained in the same sample. Movement through the gel occurs when an electric current is applied across the gel. Since the gel is immersed in buffer, the current will travel through the buffer and gel, carrying the negatively charged DNA with it toward the positive anode.

In addition to your PCR products, you will also be running a DNA Allele Ladder (shown below) that represents all the possible alleles at the BXP007 locus. This is a reference, or marker, that you can compare your PCR reactions to so you can judge their relative sizes and their identities. These are the standard sizes of all the alleles known to occur at this locus. There are 8 possible alleles, with the largest at the top of the gel and the smallest at the bottom. The sizes are, from top to bottom, 1500, 1000, 700, 500, 400, 300, 200 and 100 base pairs (bps). Allele names are based on the fragments that result from an individual DNA sample. For example, if a PCR produces 500bp and 200bp fragments, the sample has a genotype that corresponds to alleles 5 and 2 on the allele ladder. We would say that the genotype for this sample is 5-2.



DNA Gel Electrophoresis:

Materials:

- Ice bath containing PCR samples
- Allele ladder
- Loading dye
- P20 pipette
- Pipette tips
- Pre-cast 3% agarose gel
- TAE running buffer
- Gel tank
- Power source
- Light box
- Gloves – wear gloves at all times!

Protocol:

8) Set-up your gel electrophoresis tank. Be sure the wells are over the red tape in the box and that the DNA will migrate toward the positive electrode (red cable).
9) Fill your gel electrophoresis tank with running buffer.
10) Add 10µL of Orange G loading dye (from the tube labeled 'LD') to each PCR reaction tube and mix well by pipetting. **Important: use a fresh tip each time!**
11) Load 20µL of the allele ladder into the first lane. Then load 20µL of each sample into the following lanes. Each sample gets its own lane and be sure to **use a fresh tip each time.** Write down the order you load your samples in on the table below:

Lane	Sample	Load Volume
1	Allele Ladder	20µL
2		
3		
4		
5		
6		

12) Run your gel at 200V for 20 minutes. Do not let the orange dye migrate out of the gel.
13) When the gel has finished running, remove it from the electrophoresis tank and place it on the blue light box. The DNA stain that is in the agarose gel is called SybrSafe. It binds DNA by interacting with the NH group (amine) and fluoresces only upon binding. SybrSafe is excited at 502nm and emits at 540nm, so we use a blue light to visualize the DNA.
14) Take a picture to record your findings and also record the suspect genotypes under question #5.

Questions:

8) Why does DNA move through an agarose gel?

9) What are the two techniques used to create a DNA profile (used over the past two experiments)? What function does each perform?

10) What is an Allele Ladder? What is its function in DNA profiling?

11) What is required to visualize DNA following electrophoresis?

12) What are the genotypes for each of the suspects and which one matches the genotype from the crime scene?

APPENDIX C – DNA Fingerprinting questionnaire

Identification #: _____

Directions:

1. This survey has two separate sections. Each section has its own set of directions.
2. The first section of this survey contains a number of statements about science. You will be asked what you think about these statements. There are no "right" or "wrong" answers. Your opinion is what is wanted.
3. For each statement, draw a circle around the specific numeric value corresponding to how you feel about each statement. **Please circle only ONE value per statement.**

5 = Strongly Agree (SA)
4 = Agree (A)
3 = Uncertain (U)
2 = Disagree (D)
1 = Strongly Disagree (SD)

Statement	**SA**	**A**	**U**	**D**	**SD**
1. Money spent on science is well worth spending.	5	4	3	2	1
2. I would prefer to find out why something happens by doing an experiment than be being told.	5	4	3	2	1
3. I enjoy reading about things that disagree with my previous ideas.	5	4	3	2	1
4. Science is man's worst enemy.	5	4	3	2	1
5. Doing experiments is not as good as finding out information from teachers.	5	4	3	2	1
6. I dislike repeating experiments to check that I get the same results.	5	4	3	2	1
7. Public money spent on science in the last few years has been used widely.	5	4	3	2	1
8. I would prefer to do experiments rather than to read about them.	5	4	3	2	1
9. I am curious about the world in which we live.	5	4	3	2	1
10. Scientific discoveries are doing more harm than good.	5	4	3	2	1
11. I would rather agree with other people than do an experiment to find out for myself.	5	4	3	2	1
12. Finding out about new things is unimportant.	5	4	3	2	1
13. The government should spend more money on scientific research.	5	4	3	2	1

Statement	SA	A	U	D	SD
14. I would prefer to do my own experiments than to find out information from a teacher.	5	4	3	2	1
15. I like to listen to people whose opinions are different from mine.	5	4	3	2	1
16. Too many laboratories are being built at the expense of the rest of education.	5	4	3	2	1
17. I would rather find out things by asking an expert than by doing an experiment.	5	4	3	2	1
18. I find it boring to hear about new ideas.	5	4	3	2	1
19. Science helps to make life better.	5	4	3	2	1
20. I would rather solve a problem by doing an experiment than be told the answer.	5	4	3	2	1
21. In science experiments, I like to use new methods which I have not used before.	5	4	3	2	1
22. This country is spending too much money on science.	5	4	3	2	1
23. It is better to ask a teacher the answer than to find it out by doing experiments.	5	4	3	2	1
24. I am unwilling to change my ideas when evidence shows that the ideas are poor.	5	4	3	2	1
25. Science can help to make the world a better place in the future.	5	4	3	2	1
26. I would prefer to do an experiment on a topic than to read about it in science magazines.	5	4	3	2	1
27. In science experiments, I report unexpected results as well as expected ones.	5	4	3	2	1
28. Money used on scientific projects is wasted.	5	4	3	2	1
29. It is better to be told scientific facts than to find them out from experiments.	5	4	3	2	1
30. I dislike other peoples' opinions.	5	4	3	2	1

Section Two:

Directions:

1. The following questions are multiple choice (one answer) unless indicated that there are possibly multiple answers (MA).
2. Circle the answer that you think is correct and answer all the questions to the best of your ability.

1. (MA) What is the chemical composition of DNA?

protein	amino acids
nucleotides	nucleus
genes	alleles
traits	chromosomes

2. During DNA replication, what serves as the template for synthesis of a new strand?

one of the two strands of the double helix
each of the two strands of the double helix
random pieces of the both strands of the double helix

3. After DNA replication of a DNA double helix molecule, two DNA double helices result. Which of the following statements best describes the composition of the two resulting double helices?

Both strands in one of the two helices are new and both strands in the other helix are old
In each of the two helices, one of the two strands is new
Random pieces along each strand of both helices are newly synthesized

4. What are most genes made of?

protein	alleles
RNA	cells
DNA	amino acids
chromosomes	nucleus

5. (MA) In which of the following cell types within your body are genes found?

brain	heart
blood	reproductive (e.g. gametes)
eye	all of the above

6. If an organism is diploid, and has 16 chromosomes, how many sets of homologous chromosomes does it possess?

32
16
8
4
2

7. Which of the following is true of homologous chromosomes?
- They aren't usually the same size and shape
- They are inherited from different parents
- They do not pair during meiosis
- They contain different set of genes
- They contain identical DNA sequences

8. What is an allele?
- part of a gene (e.g., half of a gene)
- one of two genes in a diploid organism
- an alternate form of a gene
- a region of a chromosome where a gene is located
- a chromatid or chromosome

9. If a skin cell starts with 24 chromosomes, what number of chromosomes will it have at the end of cell division?
- 12
- 24
- 36
- 48

10. (MA) Which of the following cell types contain genetic information for the eye color of an organism?
- brain
- blood
- eye
- heart
- gamete

11. Genetic recombination refers to
- relationship between genes on the same chromosome
- independent assortment of alleles during meiosis
- the process of mutation
- co-segregation of genes during meiosis
- splicing of RNA molecules
- sorting of alleles into new combinations
- relationship between alleles of a gene

12. What is meant by the 'genetic code'?
- the differences in DNA that make individuals unique
- the specific sequence of DNA that codes for an amino acid
- the order of DNA bases within an individual
- the order of RNA bases that code for a protein

13. Which of the following molecules are the products of translation?
- DNA
- amino acids
- messenger RNAs
- cells
- chromosomes

14. Which of the following molecules are the products of transcription?
    DNA
    amino acids
    messenger RNAs
    proteins
    cells
    chromosomes

15. What is the term for the physical constitution of an organism?
    genotype
    dominant trait
    character type
    phenotype

REFERENCES

1. Gosink, M.M. and R.D. Vierstra, *Redirecting the Specificity of Ubiquitination by Modifying Ubiquitin-Conjugating Enzymes.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(20): p. 9117-9121.

2. Sullivan, M.L. and R.D. Vierstra, *Cloning of a 16-kDa ubiquitin carrier protein from wheat and Arabidopsis thaliana. Identification of functional domains by in vitro mutagenesis.* J Biol Chem, 1991. **266**(35): p. 23878-85.

3. Schneekloth, A.R., et al., *Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics.* Bioorg Med Chem Lett, 2008. **18**(22): p. 5904-8.

4. Hatakeyama, S., et al., *Targeted destruction of c-Myc by an engineered ubiquitin ligase suppresses cell transformation and tumor formation.* Cancer Res, 2005. **65**(17): p. 7874-9.

5. Melchionna, T. and A. Cattaneo, *A protein silencing switch by ligand-induced proteasome-targeting intrabodies.* J Mol Biol, 2007. **374**(3): p. 641-54.

6. Oyake, D., et al., *Targeted substrate degradation by an engineered double RING ubiquitin ligase.* Biochem Biophys Res Commun, 2002. **295**(2): p. 370-5.

7. Scheffner, M., et al., *Targeted degradation of the retinoblastoma protein by human papillomavirus E7-E6 fusion proteins.* EMBO J, 1992. **11**(7): p. 2425-31.

8. Zhou, P., et al., *Harnessing the ubiquitination machinery to target the degradation of specific cellular proteins.* Mol Cell, 2000. **6**(3): p. 751-6.

9. Krueger, K.E. and S. Srivastava, *Posttranslational protein modifications: current implications for cancer detection, prevention, and therapeutics.* Mol Cell Proteomics, 2006. **5**(10): p. 1799-810.

10. Hudson, D.F., et al., *Reverse genetics of essential genes in tissue-culture cells: 'dead cells talking'.* Trends Cell Biol, 2002. **12**(6): p. 281-7.

11. Robinson, M.S., D.A. Sahlender, and S.D. Foster, *Rapid inactivation of proteins by rapamycin-induced rerouting to mitochondria.* Dev Cell, 2010. **18**(2): p. 324-31.

12. Tamaskovic, R., et al., *Designed ankyrin repeat proteins (DARPins) from research to therapy.* Methods Enzymol, 2012. **503**: p. 101-34.

13. Fang, S. and A.M. Weissman, *A field guide to ubiquitylation.* Cell Mol Life Sci, 2004. **61**(13): p. 1546-61.

14. Ye, Y. and M. Rape, *Building ubiquitin chains: E2 enzymes at work.* Nat Rev Mol Cell

Biol, 2009. **10**(11): p. 755-64.

15. Hochstrasser, M., *Lingering mysteries of ubiquitin-chain assembly.* Cell, 2006. **124**(1): p. 27-34.

16. Komander, D. and M. Rape, *The ubiquitin code.* Annu Rev Biochem, 2012. **81**: p. 203-29.

17. Deshaies, R.J. and C.A.P. Joazeiro, *RING Domain E3 Ubiquitin Ligases.* Annual Review of Biochemistry, 2009. **78**: p. 399-434.

18. Jin, L., et al., *Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex.* Cell, 2008. **133**(4): p. 653-65.

19. Kirkin, V., et al., *A role for ubiquitin in selective autophagy.* Mol Cell, 2009. **34**(3): p. 259-69.

20. Spence, J., et al., *Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain.* Cell, 2000. **102**(1): p. 67-76.

21. Li, M., et al., *Mono- versus polyubiquitination: differential control of p53 fate by Mdm2.* Science, 2003. **302**(5652): p. 1972-5.

22. Thrower, J.S., et al., *Recognition of the polyubiquitin proteolytic signal.* EMBO J, 2000. **19**(1): p. 94-102.

23. Zhang, J., N. Zheng, and P. Zhou, *Exploring the functional complexity of cellular proteins by protein knockout.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14127-32.

24. Zhang, J. and P. Zhou, *Ectopic targeting of substrates to the ubiquitin pathway.* Methods Enzymol, 2005. **399**: p. 823-33.

25. Zhou, P., *Targeted protein degradation.* Curr Opin Chem Biol, 2005. **9**(1): p. 51-5.

26. Pickart, C.M., *Mechanisms underlying ubiquitination.* Annu Rev Biochem, 2001. **70**: p. 503-33.

27. Ma, Y.H., et al., *Targeted Degradation of KRAS by an Engineered Ubiquitin Ligase Suppresses Pancreatic Cancer Cell Growth In Vitro and In Vivo.* Molecular Cancer Therapeutics, 2013. **12**(3): p. 286-294.

28. Caussinus, E., O. Kanca, and M. Affolter, *Fluorescent fusion protein knockout mediated by anti-GFP nanobody.* Nat Struct Mol Biol, 2012. **19**(1): p. 117-21.

29. Colas, P., et al., *Targeted modification and transportation of cellular proteins.* Proc Natl Acad Sci U S A, 2000. **97**(25): p. 13720-5.

30. Lobato, M.N. and T.H. Rabbitts, *Intracellular antibodies and challenges facing their use as therapeutic agents.* Trends Mol Med, 2003. **9**(9): p. 390-6.

31. Visintin, M., M. Quondam, and A. Cattaneo, *The intracellular antibody capture technology: towards the high-throughput selection of functional intracellular antibodies for target validation.* Methods, 2004. **34**(2): p. 200-14.

32. Martineau, P., P. Jones, and G. Winter, *Expression of an antibody fragment at high levels in the bacterial cytoplasm.* J Mol Biol, 1998. **280**(1): p. 117-27.

33. Contreras-Martinez, L.M. and M.P. DeLisa, *Intracellular ribosome display via SecM translation arrest as a selection for antibodies with enhanced cytosolic stability.* J Mol Biol, 2007. **372**(2): p. 513-24.

34. Waraho, D. and M.P. DeLisa, *Versatile selection technology for intracellular protein-protein interactions mediated by a unique bacterial hitchhiker transport mechanism.* Proc Natl Acad Sci U S A, 2009. **106**(10): p. 3692-7.

35. der Maur, A.A., et al., *Direct in vivo screening of intrabody libraries constructed on a highly stable single-chain framework.* J Biol Chem, 2002. **277**(47): p. 45075-85.

36. Visintin, M., et al., *The intracellular antibody capture technology (IACT): towards a consensus sequence for intracellular antibodies.* J Mol Biol, 2002. **317**(1): p. 73-83.

37. Philibert, P., et al., *A focused antibody library for selecting scFvs expressed at high levels in the cytoplasm.* BMC Biotechnol, 2007. **7**: p. 81.

38. Pluckthun, A., *Alternative Scaffolds: Expanding the Options of Antibodies*. Recombinant antibodies for immunotherapy. 2009, New York: Cambridge University Press. xiv, 419 p.

39. Boersma, Y.L. and A. Pluckthun, *DARPins and other repeat protein scaffolds: advances in engineering and applications.* Curr Opin Biotechnol, 2011. **22**(6): p. 849-57.

40. Binz, H.K., et al., *High-affinity binders selected from designed ankyrin repeat protein libraries.* Nat Biotechnol, 2004. **22**(5): p. 575-82.

41. Pancer, Z., et al., *Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey.* Nature, 2004. **430**(6996): p. 174-80.

42. Binz, H.K., P. Amstutz, and A. Pluckthun, *Engineering novel binding proteins from nonimmunoglobulin domains.* Nat Biotechnol, 2005. **23**(10): p. 1257-68.

43. Lofblom, J., F.Y. Frejd, and S. Stahl, *Non-immunoglobulin based protein scaffolds.* Curr Opin Biotechnol, 2011. **22**(6): p. 843-8.

44. Gronwall, C. and S. Stahl, *Engineered affinity proteins--generation and applications.* J Biotechnol, 2009. **140**(3-4): p. 254-69.

45. Zhang, M., et al., *Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex.* Mol Cell, 2005. **20**(4): p. 525-38.

46. Ballinger, C.A., et al., *Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions.* Mol Cell Biol, 1999. **19**(6): p. 4535-45.

47. Rosser, M.F., et al., *Chaperone functions of the E3 ubiquitin ligase CHIP.* J Biol Chem, 2007. **282**(31): p. 22267-77.

48. Qian, S.B., et al., *Engineering a ubiquitin ligase reveals conformational flexibility required for ubiquitin transfer.* J Biol Chem, 2009. **284**(39): p. 26797-802.

49. Jeong, H., et al., *Genome sequences of Escherichia coli B strains REL606 and BL21(DE3).* J Mol Biol, 2009. **394**(4): p. 644-52.

50. Xu, Z., et al., *Structure and interactions of the helical and U-box domains of CHIP, the C terminus of HSP70 interacting protein.* Biochemistry, 2006. **45**(15): p. 4749-59.

51. Koch, H., et al., *Direct selection of antibodies from complex libraries with the protein fragment complementation assay.* J Mol Biol, 2006. **357**(2): p. 427-41.

52. Juers, D.H., et al., *High resolution refinement of beta-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for alpha-complementation.* Protein Sci, 2000. **9**(9): p. 1685-99.

53. Pickart, C.M., *Targeting of substrates to the 26S proteasome.* FASEB J, 1997. **11**(13): p. 1055-66.

54. Kundrat, L. and L. Regan, *Identification of residues on Hsp70 and Hsp90 ubiquitinated by the cochaperone CHIP.* J Mol Biol, 2010. **395**(3): p. 587-94.

55. Xu, P., et al., *Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation.* Cell, 2009. **137**(1): p. 133-45.

56. Saeki, Y., et al., *Lysine 63-linked polyubiquitin chain may serve as a targeting signal for the 26S proteasome.* EMBO J, 2009. **28**(4): p. 359-71.

57. Arndt, V., et al., *Chaperone-assisted selective autophagy is essential for muscle maintenance.* Curr Biol, 2010. **20**(2): p. 143-8.

58. Kraft, C., M. Peter, and K. Hofmann, *Selective autophagy: ubiquitin-mediated recognition and beyond.* Nat Cell Biol, 2010. **12**(9): p. 836-41.

59. Xu, Z., et al., *Interactions between the quality control ubiquitin ligase CHIP and ubiquitin conjugating enzymes.* BMC Struct Biol, 2008. **8**: p. 26.

60. Fujiwara, K., et al., *A single-chain antibody/epitope system for functional analysis of protein-protein interactions.* Biochemistry, 2002. **41**(42): p. 12729-38.

61. Krebber, C., et al., *Co-selection of cognate antibody-antigen pairs by selectively-infective*

phages. FEBS Lett, 1995. **377**(2): p. 227-31.

62. Koide, A., et al., *High-affinity single-domain binding proteins with a binary-code interface.* Proc Natl Acad Sci U S A, 2007. **104**(16): p. 6632-7.

63. Gilbreth, R.N., et al., *A dominant conformational role for amino acid diversity in minimalist protein-protein interfaces.* J Mol Biol, 2008. **381**(2): p. 407-18.

64. Yang, F., et al., *Novel fold and capsid-binding properties of the lambda-phage display platform protein gpD.* Nat Struct Biol, 2000. **7**(3): p. 230-7.

65. Shaw, D., et al., *The crystal structure of JNK2 reveals conformational flexibility in the MAP kinase insert and indicates its involvement in the regulation of catalytic activity.* J Mol Biol, 2008. **383**(4): p. 885-93.

66. Karlsson, A.J., et al., *Engineering antibody fitness and function using membrane-anchored display of correctly folded proteins.* J Mol Biol, 2012. **416**(1): p. 94-107.

67. Qian, S.B., et al., *CHIP-mediated stress recovery by sequential ubiquitination of substrates and Hsp70.* Nature, 2006. **440**(7083): p. 551-5.

68. Stankiewicz, M., et al., *CHIP participates in protein triage decisions by preferentially ubiquitinating Hsp70-bound substrates.* FEBS J, 2010. **277**(16): p. 3353-67.

69. Smith, M.C., et al., *The E3 ubiquitin ligase CHIP and the molecular chaperone Hsc70 form a dynamic, tethered complex.* Biochemistry, 2013. **52**(32): p. 5354-64.

70. Swinney, D.C., et al., *Bi-substrate kinetic analysis of an E3-ligase-dependent ubiquitylation reaction.* Methods Enzymol, 2005. **399**: p. 323-33.

71. Fushman, D. and K.D. Wilkinson, *Structure and recognition of polyubiquitin chains of different lengths and linkage.* F1000 Biol Rep, 2011. **3**: p. 26.

72. Purbeck, C., Z.M. Eletr, and B. Kuhlman, *Kinetics of the transfer of ubiquitin from UbcH7 to E6AP.* Biochemistry, 2010. **49**(7): p. 1361-3.

73. Kleiger, G., et al., *Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates.* Cell, 2009. **139**(5): p. 957-68.

74. Jiang, J., et al., *CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation.* J Biol Chem, 2001. **276**(46): p. 42938-44.

75. Cyr, D.M., J. Hohfeld, and C. Patterson, *Protein quality control: U-box-containing E3 ubiquitin ligases join the fold.* Trends Biochem Sci, 2002. **27**(7): p. 368-75.

76. Kampinga, H.H., et al., *Overexpression of the cochaperone CHIP enhances Hsp70-dependent folding activity in mammalian cells.* Mol Cell Biol, 2003. **23**(14): p. 4948-58.

77. Meacham, G.C., et al., *The Hsc70 co-chaperone CHIP targets immature CFTR for*

proteasomal degradation. Nat Cell Biol, 2001. **3**(1): p. 100-5.

78. Connell, P., et al., *The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins.* Nat Cell Biol, 2001. **3**(1): p. 93-6.

79. Kumar, P., et al., *CHIP and HSPs interact with beta-APP in a proteasome-dependent manner and influence Abeta metabolism.* Hum Mol Genet, 2007. **16**(7): p. 848-64.

80. Hwang, J.R., C. Zhang, and C. Patterson, *C-terminus of heat shock protein 70-interacting protein facilitates degradation of apoptosis signal-regulating kinase 1 and inhibits apoptosis signal-regulating kinase 1-dependent apoptosis.* Cell Stress Chaperones, 2005. **10**(2): p. 147-56.

81. Shang, Y., et al., *CHIP functions an E3 ubiquitin ligase of Runx1.* Biochem Biophys Res Commun, 2009. **386**(1): p. 242-6.

82. Wang, L., et al., *Molecular mechanism of the negative regulation of Smad1/5 protein by carboxyl terminus of Hsc70-interacting protein (CHIP).* J Biol Chem, 2011. **286**(18): p. 15883-94.

83. Nikolay, R., et al., *Dimerization of the human E3 ligase CHIP via a coiled-coil domain is essential for its activity.* J Biol Chem, 2004. **279**(4): p. 2673-8.

84. Sha, Y., et al., *A critical role for CHIP in the aggresome pathway.* Mol Cell Biol, 2009. **29**(1): p. 116-28.

85. Lorenz, S., et al., *Macromolecular juggling by ubiquitylation enzymes.* BMC Biol, 2013. **11**: p. 65.

86. Barkhordarian, H., et al., *Isolating recombinant antibodies against specific protein morphologies using atomic force microscopy and phage display technologies.* Protein Eng Des Sel, 2006. **19**(11): p. 497-502.

87. Lecerf, J.M., et al., *Human single-chain Fv intrabodies counteract in situ huntingtin aggregation in cellular models of Huntington's disease.* Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4764-9.

88. Lynch, S.M., C. Zhou, and A. Messer, *An scFv intrabody against the nonamyloid component of alpha-synuclein reduces intracellular aggregation and toxicity.* J Mol Biol, 2008. **377**(1): p. 136-47.

89. Thurston, T.L., et al., *The TBK1 adaptor and autophagy receptor NDP52 restricts the proliferation of ubiquitin-coated bacteria.* Nat Immunol, 2009. **10**(11): p. 1215-21.

90. Zhang, X. and S.B. Qian, *Chaperone-mediated hierarchical control in targeting misfolded proteins to aggresomes.* Mol Biol Cell, 2011. **22**(18): p. 3277-88.

91. Gamerdinger, M., et al., *BAG3 mediates chaperone-based aggresome-targeting and*

selective autophagy of misfolded proteins. EMBO Rep, 2011. **12**(2): p. 149-56.

92. Wong, E.S., et al., *Autophagy-mediated clearance of aggresomes is not a universal phenomenon.* Hum Mol Genet, 2008. **17**(16): p. 2570-82.

93. Shin, Y., et al., *The co-chaperone carboxyl terminus of Hsp70-interacting protein (CHIP) mediates alpha-synuclein degradation decisions between proteasomal and lysosomal pathways.* J Biol Chem, 2005. **280**(25): p. 23727-34.

94. Schrader, E.K., K.G. Harstad, and A. Matouschek, *Targeting proteins for degradation.* Nat Chem Biol, 2009. **5**(11): p. 815-22.

95. Tsai, Y.C., et al., *Targeting botulinum neurotoxin persistence by the ubiquitin-proteasome system.* Proc Natl Acad Sci U S A, 2010. **107**(38): p. 16554-9.

96. Rotin, D. and S. Kumar, *Physiological functions of the HECT family of ubiquitin ligases.* Nat Rev Mol Cell Biol, 2009. **10**(6): p. 398-409.

97. Kamadurai, H.B., et al., *Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3.* Elife, 2013. **2**: p. e00828.

98. Huang, L., et al., *Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade.* Science, 1999. **286**(5443): p. 1321-6.

99. Ogunjimi, A.A., et al., *Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain.* Mol Cell, 2005. **19**(3): p. 297-308.

100. Lin, D.Y., J. Diao, and J. Chen, *Crystal structures of two bacterial HECT-like E3 ligases in complex with a human E2 reveal atomic details of pathogen-host interactions.* Proc Natl Acad Sci U S A, 2012. **109**(6): p. 1925-30.

101. Hicks, S.W. and J.E. Galan, *Hijacking the host ubiquitin pathway: structural strategies of bacterial E3 ubiquitin ligases.* Curr Opin Microbiol, 2010. **13**(1): p. 41-6.

102. Zhu, Y., et al., *Structure of a Shigella effector reveals a new class of ubiquitin ligases.* Nat Struct Mol Biol, 2008. **15**(12): p. 1302-8.

103. Turner, G.C., F. Du, and A. Varshavsky, *Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway.* Nature, 2000. **405**(6786): p. 579-83.

104. Ito, S., Y.H. Song, and T. Imaizumi, *LOV domain-containing F-box proteins: light-dependent protein degradation modules in Arabidopsis.* Mol Plant, 2012. **5**(3): p. 573-82.

105. Yoshida, Y., et al., *E3 ubiquitin ligase that recognizes sugar chains.* Nature, 2002. **418**(6896): p. 438-42.

106. Pickart, C.M., *Back to the future with ubiquitin.* Cell, 2004. **116**(2): p. 181-90.

107. Parizek, P., et al., *Designed ankyrin repeat proteins (DARPins) as novel isoform-specific*

intracellular inhibitors of c-Jun N-terminal kinases. ACS Chem Biol, 2012. **7**(8): p. 1356-66.

108. Kummer, L., et al., *Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries.* Proc Natl Acad Sci U S A, 2012. **109**(34): p. E2248-57.

109. Coffey, R., *Outsmarting Dengue Fever by Vaccinating Mosquitoes.* Scientific American, 2011(April 2011).

110. Elrod, S., *Genetic Concepts Inventory.* 2007.

111. Fraser, B.J., *Test of Science-Related Attitudes Handbook,* A.C.f.E. Research, Editor. 1981, Radfor House.